

1441 TTCTCTTCAACGGTTCTGTCAATTCAGGACCAGGATTCA 1480  
 5  
 1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520  
 10  
 1521 CATTGAGAATAGAGGGTATATTGAAGTTCCAATTCACCTC 1560  
 1561 CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG 1600  
 15  
 1601 CTTCTGTGACCCCTATTACCTCAACGTTAATTGGGGTAA 1640  
 1641 TTCATCCATCTTCTCCAATACAGTTCAGCTACAGCTACC 1680  
 20  
 1681 TCCTTGGATAATCTCCAATCCAGCGATTCGGTTACTTTG 1720  
 1721 AAAGTGCCAATGCTTTTACATCTTCACTCGGTAACATCGT 1760  
 25  
 1761 GGGTGTTAGAACTTTAGTGGGACTGCAGGAGTGATTATC 1800  
 1801 GACAGATTGAGATTCAATCCAGTTACTGCAACACTCGAGG 1840  
 30  
 1841 CTGAGTACAACCTTGAGAGAGCCGAGAAGGCTGTGAACGC 1880  
 35  
 1881 CCTCTTTACCTCCACCAATCAGCTTGGCTTGAAAACATAAC 1920  
 1921 GTTACTGACTATCACATTGACCAAGTGTCCAACCTTGGTCA 1960  
 40  
 1961 CCTACCTTAGCGATGAGTTCTGCCTCGACGAGAAGCGTGA 2000  
 2001 ACTCTCCGAGAAAAGTTAAACACGCCAAGCGTCTCAGCGAC 2040  
 45  
 2041 GAGAGGAATCTCTTGCAAGACTCCAACCTTCAAGACATCA 2080  
 2081 ACAGGCAGCCAGAACGTGGTTGGGGTGGAAGCACCGGGGAT 2120  
 50  
 2121 CACCATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTAC 2160  
 55  
 2161 GTCACCCTCTCCGGAACTTTCGACGAGTGCTACCCCTACCT 2200

5	2201	ACTTGTACCAGAAGATCGATGAGTCCAAACTCAAAGCCTT	2240
10	2241	CACCAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAA	2280
15	2281	GACCTTGAATCTACTCGATCAGGTACAATGCCAAGCAGC	2320
20	2321	AGACCGTGAATGTCCCAGGTACTGGTTCCCTCTGGCCACT	2360
25	2361	TTCTGCCCAATCTCCCATTTGGGAAGTGTGGAGAGCCTAAC	2400
30	2401	AGATGCGCTCCACACCTTGAGTGGAACTCCTGACTTGGACT	2440
35	2441	GCTCCTGCAGGGATGGCGAGAAGTGTGCCACCATTCTCA	2480
40	2481	TCACCTTCTCCTTGGACATCGATGTGGGATGTACTGACCTG	2520
45	2521	AATGAGGACCTCGGAGTCTGGGTTCATCTTCAAGATCAAGA	2560
50	2561	CCCAAGACGGACACGCAAGACTTGGCAACCTTGAGTTTCT	2600
55	2601	CGAAGAGAAACCATTTGGTCGGTGAAGCTCTCGCTCGTGTG	2640
60	2641	AAGAGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAAAC	2680
65	2681	TCGAATGGGAAACTAACATCGTTTACAAGSAGGCCAAAGA	2720
70	2721	GTCCGTGGATGCTTTGTTCGTGAACCTCCCAATATGATCAG	2760
75	2761	TTGCAAGCCGACACCAACATCGCCATGATCCACGCCGCGAG	2800
80	2801	ACAAACGTGTGCACAGCATTCGTGAGGCTTACTTGCCTGA	2840
85	2841	GTTGTCCGTGATCCCTGCTGGAACGCTGCCATCTTCGAG	2880
90	2881	GAACTTGAGGGACGTATCTTTACCGCATTCCTCTGTACG	2920
95	2921	ATGCCAGAAACGTCATCAAGAACGGTGACTTCAACAATGG	2960

```

      .       .       .       .
2961 CCTCAGCTGCTGGAATGTGAAAGGTCATGTGGACGTGGAG 3000
      .       .       .       .
      .       .       .       .
3001 GAACAGAACCAATCAGCGTCCCGTCCTGGTTGTGCCTGAGT 3040
      .       .       .       .
      .       .       .       .
10 3041 GGGGAAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGG 3080
      .       .       .       .
      .       .       .       .
3081 TAGAGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGA 3120
      .       .       .       .
      .       .       .       .
15 3121 TACGGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACA 3160
      .       .       .       .
      .       .       .       .
3161 ACACCGACGAGCTTAAGTTCTCCAACGCGTCGAGGAAGA 3200
      .       .       .       .
      .       .       .       .
20 3201 AATCTATCCCAACAACACCGTTACTTGCAACGACTACACT 3240
      .       .       .       .
      .       .       .       .
25 3241 GTGAATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTA 3280
      .       .       .       .
      .       .       .       .
3281 ACAGAGGTTACAACGAAGCTCCTTCCGTTCTGTGCTGACTA 3320
      .       .       .       .
      .       .       .       .
30 3321 TGCCCTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGA 3360
      .       .       .       .
      .       .       .       .
3361 CGTGAGAACCCTTGCCGAGTTCAACAGAGGTTACAGGGACT 3400
      .       .       .       .
      .       .       .       .
35 3401 ACACACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGA 3440
      .       .       .       .
      .       .       .       .
40 3441 GTACTTTCCTGAGACCGACAAAAGTGTGGATCGAGATCGGT 3480
      .       .       .       .
      .       .       .       .
3481 GAAACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTC 3520
      .       .       .       .
      .       .       .       .
45 3521 TCTTGATGGAGGAA 3534,

```

H. einem Struktur-Gen, das für ein insektizides Protein von *B.t.t.* codiert, mit der Sequenz:

```

      .       .       .       .
1  ATGACTGCAGACAACAACCGAAGCCCTCGACAGTTCTA 40
      .       .       .       .
      .       .       .       .
55 41  CCACTAAGGATGTTATCCAGAAGGGTATCTCCGTTGTGGG 80

```

81 AGACCTCTTGGGCGTGGTTGGATTCCCTTCGGTGGAGCC 120  
 5  
 121 CTCGTGAGCTTCTATACAACTTTCTCAACACCATTTGGC 160  
 161 CAAGCGAGGACCCTTGGAAAGCATTATGGAGCAAGTTGA 200  
 10  
 201 AGCTCTTATGGATCAGAAGATTGCAGATTATGCCAAGAAC 240  
 241 AAGGCTTTGGCAGAACTCCAGGGCCCTTCAGAACAAATGTGG 280  
 15  
 281 AGGACTACGTGAGTGCATTGTCCAGCTGGCAGAAGAACC 320  
 321 TGTTAGCTCCAGAAATCCTCACAGCCAAGGTAGGATCAGA 360  
 20  
 361 GAGTTGTTCTCTCAAGCCGAATCCCACTTCAGAAATTCCA 400  
 25  
 401 TGCCTAGCTTTGCTATCTCCGGTTACGAGGTTCTTTTCCT 440  
 441 CACTACCTATGCTCAGCTGCCAACACCCACTTGTTTCTC 480  
 30  
 481 CTTAAGGACGCTCAAATCTATGGAGAAGAGTGGGGATACG 520  
 521 AGAAAGAGGACATTGCTGAGTTCTACAAGCGTCAACTTAA 560  
 35  
 561 GCTCACCCAAAGAGTACACTGACCATTGCGTGAAATGGTAT 600  
 40  
 601 AACGTTGGTCTCGATAAGCTCAGAGGCTCTTTCCTACGAGT 640  
 641 CTTGGGTGAAC TTCAACAGATACAGGAGAGAGATGACCTT 680  
 45  
 681 GACTGTGCTCGATCTTATCGCACTCTTTCCCTTGATACGAT 720  
 721 GTGAGACTCTACCCAAAGGAAGTGAAACTGAGCTTACCA 760  
 50  
 761 GAGACGTGCTCACTGACCTATTGTGCGAGTCAACAACCT 800  
 55  
 801 TAGGGGTTATGGAAC TACCTTCAGCAATATCGAAAAC TAC 840

	841	ATTAGGAAACCACATCTCTTCGACTATCTTCACAGAATTC	880
5	881	AATTCCACACAAGGTTTCAACCAGGATACTATGGTAACGA	920
10	921	CTCCTTCAACTATTGGTCCGGTAAGTATGTTTCCACCAGA	960
	961	CCAAGCATTGGATCTAATGACATCATCACATCTCCCTTCT	1000
15	1001	ATGGTAACAAGTCCAGTGAACCTGTGCAGAACCTTGAGTT	1040
	1041	CAACGGCGAGAAAGTCTATAGAGCCGTCGCAAACACCAAT	1080
20	1081	CTCGCTGTGTGGCCATCCGCAGTTTACTCAGGCGTCACAA	1120
	1121	AGGTGGAGTTTAGTCAGTATAACGATCAGACCGATGAGGC	1160
25	1161	CAGCACCCAGACTTACGACTCCAAACGTAACGTTGGCGCA	1200
30	1201	GTCTCTTGGGATTCTATCGACCAATTGCCTCCAGAAACCA	1240
	1241	CAGACGAACCATTTGGAGAAGGGCTACAGCCACCAACTTAA	1280
35	1281	CTATGTGATGTGCTTCTTGATGCAAGGTTCCAGAGGGACC	1320
	1321	ATTCCAGTGTGACCTGGACACACAAGTCCGTGGACTTCT	1360
40	1361	TCAACATGATCGATAGCAAGAAGATCACTCAACTTCCCTT	1400
	1401	GGTGAAAGCCTACAAGCTGCAATCTGGTGCTTCCGTTGTC	1440
45	1441	GCAGGTCCCAGATTCACTGGAGGTGACATCATCCAGTGCA	1480
50	1481	CAGAGAACGGCAGCGCAGCTACTATCTACGTGACACCTGA	1520
	1521	TGTGTCTTACTCTCAGAAGTACAGGGCAGGTATTCAATTAC	1560
55	1561	GCATCTACCAGCCAGATCACCTTCACACTCAGCTTGGATG	1600

1601 GAGCACCCCTTCAACCAGTATTACTTTGACAAGACCATCAA 1640  
 1641 CAAAGGTGACACTCTCACATACAATAGCTTCAACTTGGCA 1680  
 1681 AGTTTCAGCACACCATTGAACTCTCAGGCAACAATCTTC 1720  
 1721 AGATCGGCGTCACCGGTCTCAGCGCCGGAGACAAAGTCTA 1760  
 1761 CATCGACAAGATTGAGTTCATCCCAGTGAAC 1791,

I. einem Struktur-Gen, das für ein insektizides Protein von *B. t. entomocidus* codiert, mit der Sequenz:

1 ATGGAGGAGAACAACCAAAACCAATGCATTCCATACAAC 40  
 41 GCTTGAGTAACCCAGAAGAGGTATTGCTTGATGGAGAACG 80  
 81 CATTTCACCGGTAACCTCTTCATCGACATCTCCTTGTC 120  
 121 TTGGTCCAGTTTCTGGTCAGCAACTTCGTGCCAGGTGGTG 160  
 161 GGTTCCCTTGTCGGACTAATTGACTTCGTCTGGGGTATCGT 200  
 201 TGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATTGAG 240  
 241 CAGTTGATCAACGAGAGGATCGCTGAGTTGCCAGGAACG 280  
 281 CTGCCATCGCTAACTTGAAGGATTGGGCAATAACTTCAA 320  
 321 CATCTATGTGGAGGCCCTTCAAAGAGTGGGAAGAGGACCT 360  
 361 AACCAACCCAGAGACCCGCACTAGGGTGATCGACAGATTCA 400  
 401 GAATCTTGGACGGCCTCTTGGAGAGAGATATCCCATCCTT 440  
 441 CAGAACTCTCTGGCTTCGAAGTTCCTCTCTTGTCGGTGTAC 480

481 GCTCAAGCAGCTAATCTTCACCTCGCTATCCTTCGAGACA 520  
 521 GTGTCATCTTTGGGGAAAGGTGGGGATTGACCACTATCAA 560  
 561 CGTCAATGAGAATTACAACAGACTTATCAGGCACATTGAC 600  
 601 GAGTACGCCGACCACTGTGCTAACACCTACAACCGTGGCT 640  
 641 TGAACAATCTCCCTAAGTCTACTTATCAAGATTGGATTAC 680  
 681 CTACAACAGGTTGAGGAGAGACTTGACCCCTCACAGTTTTG 720  
 721 GACATTGCAGCTTTCTTCCCGAACTATGACAACAGGAGAT 760  
 761 ACCCTATCCAACCAAGTGGGTCAACTTACCAGAGAACTCTA 800  
 801 TACTGACCCACTTATCAACTTCAACCCCTCAGTTGCAAAGT 840  
 841 GTCGCCCAACTTCCCACATTCAACGTCATGGAGTCCAGCC 880  
 881 GTATCAGGAACCCACACTTGTTTGACATCTTGAACAACCT 920  
 921 TACTATCTTCACCGATTGGTTTCAGCGTTGGGCGTAACTTC 960  
 961 TATTGGGGTGGACACAGGGTCATCTCCTCTCTTATTGGAG 1000  
 1001 GTGGGAACATTACCTCTCCTATCTATGGACGTGAGGCAAA 1040  
 1041 CCAGGAGCCACCACGTAGTTTCACCTTCAACGGTCCAGTC 1080  
 1081 TTCAGAACCTTGTCTAACCCCTACCTTGAGATTGCTCCAGC 1120  
 1121 AACCTTGGCCAGCTCCACCTTTCAACCTTAGAGGTGTTGA 1160  
 1161 GGGCGTTGAGTTCTCTACTCCTACCAACTCCTTCACTTAC 1200  
 1201 AGAGGTAGAGGAACCGTTGATTCTTGGACCGAACTCCAC 1240

5	1241	CAGAGGACAATAGCGTGCCACCCAGGGAAGGCTACTCCCA	1280
	1281	CAGGTTGTGCCACGCAACCTTCGTGCAGCGTTCGGAACT	1320
10	1321	CCATTCCCTCACTACAGGAGTTGTGTTCTCATGGACTGATC	1360
	1361	GTAGTGCTACTCTCACTAATACCATTGATCCCAGAGGGAT	1400
15	1401	CAATCAAATCCCATTGGTCAAGGGTTTCGGTGTGTGGGGA	1440
	1441	GGAACCTTCTGTCAATCACAGGACCAGGCTTCACAGGAGGTG	1480
20	1481	ATATTCTTAGAAGAAACACTTTTGGCGACTTTGTGAGCCT	1520
	1521	CCAAAGTTAACATCAACTCTCCAATTACTCAAAGATATCGT	1560
25	1561	CTCAGGTTTCGTTACGCATCTTCCCGTGACGCTAGAGTCA	1600
	1601	TCGTGCTCACCGGAGCAGCTTCTACCGGTGTCGGTGGACA	1640
30	1641	AGTCTCCGTGAACATGCCACTCCAGAAGACTATGGAGATC	1680
	1681	GGCGAGAAGCTTGACATCCAGGACCTTCAGATACACCGACT	1720
35	1721	TCTCTAACCCTTTCAGTTTCCGTGCCAACCCCTGACATCAT	1760
40	1761	TGGCATTAGCGAACAACCTCTCTTTGGAGCTGGTAGCATC	1800
	1801	TCATCTGGCGAATTGTACATTGACAAGATTGAGATCATTC	1840
45	1841	TTGCCGACGCTACCTTCGAGGCTGAGTCTGACCTTGAGAG	1880
	1881	AGCCCAGAAGGCTGTGAACGCCCTCTTTACCTCCTCTAAT	1920
50	1921	CAGATTGGCTTGAAAACCTGACGTTACTGACTATCACATTG	1960
55	1961	ACCAAGTGTCCAACCTTGGTCGACTGCCTTAGCGATGAGTT	2000



2001	CTGCCTCGACGAGAAGCGTGAACCTCTCCGAGAAAGTTAAA	2040
5		
2041	CACGCCAAGCGTCTCAGCGACGAGAGGAATCTCTTGCAAG	2080
10		
2081	ACCCCAACTTCAGAGGCATCAACAGGCAGCCAGACCGTGG	2120
2121	TTGGAGAGGAAGCACCACATCACCATCCAAGGAGGCGAC	2160
15		
2161	GATGTGTTCAAGGAGAACTACGTCACCCTCCCAGGAACTG	2200
2201	TGACGAGTGCTACCTTACCTACTTGTACCAGAAGATCGA	2240
20		
2241	TGAGTCCAAACTCAAAGCCTACACCAGGTATGAACTTAGA	2280
2281	GGCTACATCGAAGACAGCCAAGACCTTGAAATCTACCTCA	2320
25		
2321	TCAGGTACAATGCCAAGCACGAGATCGTGAATGTCCCAGG	2360
2361	TACTGGTTCCCTCTGGCCACTTTCTGCCCCAAATGCCCAAT	2400
30		
2401	GGGAAGTGTGGAGAGCCTAACAGATGCGCTCCACACCTTG	2440
2441	AGTGGAAATCCTGACTTGGACTGCTCCTGCAGGGATGSCGA	2480
35		
2481	GAAGTGTGCCACCATTCTCATCACTTCACCTTGGACATC	2520
40		
2521	GATGTGGGATGTACTGACCTGAATGAGGACCTCGGAGTCT	2560
2561	GGGTCACTCTTCAAGATCAAGACCCAAGACGGACACGCAAG	2600
45		
2601	ACTTGGCAACCTTGAGTTTCTCGAAGAGAAACCATTGCTC	2640
2641	GGTGAAGCTCTCGCTCGTGTGAAGAGAGCAGAGAAGAAGT	2680
50		
2681	GGAGGGACAACSTGAGAACTCCAACCTCGAGACTAACAT	2720
2721	CGTTTACAAGGAGGCCAAAGAGTCCGTGGATGCTTTGTTC	2760
55		

	2761	GTGAACTCCCAATATGATAGGTTGCAAGTGGACACCAACA	2800
5			
	2801	TCGCCATGATCCACGCTGCAGACAAACGTGTGCACAGGAT	2840
10			
	2841	TCGTGAGGCTTACTTGCCTGAGTTGTCCGTGATCCCTGGT	2880
	2881	GTGAACGCTGCCATCTTCGAGGAACTTGAGGGACGTATCT	2920
15			
	2921	TTACCGCATACTCCTTGTACGATGCCAGAAACGTCATCAA	2960
	2961	GAACGGTGACTTCAACAATGGCCTCTTGTGCTGGAATGTG	3000
20			
	3001	AAAGGTCATGTGGACGTGGAGGAACAGAACATCACCGTT	3040
	3041	CCGTCTCGGTTATCCCTGAGTGGGAAGCTGAAGTGTCCCA	3080
25			
	3081	AGAGGTTAGAGTCTGTCCAGGTAGAGGGCTACATTCTCCGT	3120
	3121	GTGACCGCTTACAAGGAGGGATACGGTGAGGGTTGCGTGA	3160
30			
	3161	CCATCCACGAGATCGAGGACAACCCGACGAGCTTAAGTT	3200
	3201	CTCCAACTGCGTTCGAGGAAGAAGTCTATCCCAACAACACC	3240
35			
	3241	GTTACTTGCAACAACCTACACTGGGACCCAGGAAGAGTACG	3280
40			
	3281	AAGGTACCTACACTAGCCGTAACCAAGGTTACGACGAAGC	3320
	3321	TTACGGAAACAATCCTTCCGTTCTGCTGACTATGCCTCC	3360
45			
	3361	GTGTACGAGGAGAAAATCCTACACAGATGGCAGACGTGAGA	3400
	3401	ACCCTTGCAGTCCACAGAGGTTACGGTGACTACACACC	3440
50			
	3441	ACTTCCAGCAGGCTATGTTACCAAGGACCTTGAGTACTTT	3480
55			
	3481	CCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAAACCG	3520

3521 AGGGAACCTTCATCGTGGACAGCGTGGAGCTTCTCTTGAT 3560

3561 GGAGGAA 3567.

J. einem Struktur-Gen, das für ein insektizides P2-Protein codiert, mit der Sequenz:

```

1  ATGGACAACAACGCTCTTGAACCTCTGGTAGAACAACCATCT  40
41 GCGACGCATACAACGTCGTGGCTCACGATCCATTTCAGCTT  80
81 CGAACACAAGAGCCTCGACACTATTGAGAAGGAGTGGATG  120
121 GAATGGAAACGTACTGACCACTCTCTCTACGTCGCACCTG  160
161 TGGTTGGAACAGTGTCCAGCTTCCTTCTCAAGAAGGTCGG  200
201 CTCTCTCATCGGAAAACGTATCTTGTCCGAACCTCTGGGGT  240
241 ATCATCTTTCCATCTGGGTCCACTAATCTCATGCAAGACA  280
281 TCTTGAGGGAGACCGAACAGTTTCTCAACCAGCSTCTCAA  320
321 CACTGATACCTTGGCTAGAGTCAACGCTGAGTTGATCGGT  360
361 CTCCAAGCAAACATTTCGTGAGTTCAACCAGCAAGTGGACA  400
401 ACTTCTTGAATCCAACTCAGAATCCTGTGCCTCTTTCCAT  440
441 CACTTCTTCCGTGAACACTATGCAGCAACTCTTCCTCAAC  480
481 AGATTGCCTCAGTTTCAGATTCAAGGCTACCAAGTTGCTCC  520
521 TTCTTCCACTCTTTGCTCAGGCTGCCAACATGCACTTGTC  560
561 CTTTCATACGTGACGTGATCCTCAACGCTGACGAATGGGGA  600

```

5 601 ATCTCTGCAGCCACTCTTAGGACATACAGAGACTACTTGA 640  
 641 GGAACTACACTCGTGATTACTCCAACATATGCATCAACAC 680  
 10 681 TTATCAGACTGCCCTTCGTGGACTCAATACTAGGCTTCAC 720  
 721 GACATGCTTGAGTTCAGGACCTACATGTTCCCTTAACGTGT 760  
 15 761 TTGAGTACGTCAGCATTGGAGTCTCTTCAAGTACCAGAG 800  
 801 CTTGATGGTGTCTCTGGAGCCAATCTCTACGCCTCTGGC 840  
 20 841 AGTGGACCACAGCAAACCTCAGAGCTTCACAGCTCAGAACT 880  
 881 GGCCATTCTTGTATAGCTTGTCCAAGTCAACTCCAACCTA 920  
 25 921 CATTCTCAGTGTATCTCTGGGACCAGACTCTCCATAACC 960  
 961 TTTCCCAACATTGGTGGACTTCCAGGCTCCACTACAACCC 1000  
 1001 ATAGCCTTAACTCTGCCAGAGTGAACCTACAGTGGAGGTGT 1040  
 35 1041 CAGCTCTGGATTGATTGGTGCAACTAACTTGAACCACAAC 1080  
 1081 TTCAATTGCTCCACCGTCTTGCCACCTCTGAGCACACCGT 1120  
 40 1121 TTGTGAGGTCTCTGGCTTGACAGCGGTACTGATCGCGAAGG 1160  
 1161 AGTTGCTACCTCTACAAACTGGCAAACCGAGTCTCTCCAA 1200  
 45 1201 ACCACTCTTAGCCTTCGGTGTGGAGCTTTCTCTGCACGTG 1240  
 1241 GGAATTCAAACTACTTTCCAGACTACTTCATTAGGAACAT 1280  
 50 1281 CTCTGGTGTTCCTCTCGTCATCAGGAATGAAGACCTCACC 1320  
 55 1321 CGTCCACTTCATTACAACCAGATTAGGAACATCGAGTCTC 1360

1361 CATCCGGTACTCCAGGAGGTGCAAGAGCTTACCTCGTGTC 1400  
 1401 TGTCCATAACAGGAAGAACAACATCTACGCTGCCAACGAG 1440  
 1441 AATGGCACCATGATTACCTTGCAACAGAAGATTACACTG 1480  
 1481 GATTACCATCTCTCCAATCCATGCTACCCAAGTGAACAA 1520  
 1521 TCAGACACGCACCTTCATCTCCGAAAAGTTCGGAAATCAA 1560  
 1561 GGTGACTCCTTGAGGTTGAGCAATCCAACACTACCGCTA 1600  
 1601 GGTACACTTTGAGAGGCAATGGAAACAGCTACAACCTTTA 1640  
 1641 CTTGAGAGTTAGCTCCATTGGTAACCTCCACCTCCGTGTT 1680  
 1681 ACCATCAACGGACGTGTTTACACAGTCTCTAATGTGAACA 1720  
 1721 CTACAACGAACAATGATGGCGTTAACGACAACGGAGCCAG 1760  
 1761 ATTCAGCGACATCAACATTGGCAACATCGTGGCCTCTGAC 1800  
 1801 AACACTAACGTTTACTTTGGACATCAATGTGACCCCTCAATT 1840  
 1841 CTGGAACCTCCATTGATCTCATGAACATCATGTTTGTGCC 1880  
 1881 AACTAACCTCCCTCCATTGTAC 1902

oder

K. einer Struktur-Gen-Sequenz, die für ein Fusionsprotein codiert, das die N-terminalen 610 Aminosäuren von  
 B.t.k. HD-1 und die C-terminalen 567 Aminosäuren von B.t.k. HD-73 aufweist, welches Gen die Sequenz hat:

1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40

	41	ACTGCTTGAGTAACCCAGAGTTGAAGTACTTGGTGGAGA	80
5	81	ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG	120
10	121	TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG	160
15	161	CTGGGTTCGTTCTCGGACTAGTTGACATCATCTSGGGTAT	200
20	201	CTTTGGTCCATCTCAATGGGATGCATTCTGTGTGCAAATT	240
25	241	GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA	280
30	281	ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA	320
35	321	CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT	360
40	361	CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT	400
45	401	TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT	440
50	441	GTTGCGAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTS	480
55	481	TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCCTCGAG	520
	521	ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTTCGATGCTGC	560
	561	AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT	600
	601	GGAAACTACACCGACCACGCTGTTGTTGGTACAACACTG	640
	641	GCTTGGAGCGTGCTCGGGTCTCTGATTCTAGAGATTGGAT	680
	681	TAGATACAACCAAGTTCAGGAGAGAATTGACCCCTCACAGTT	720
	721	TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA	760
	761	CCTACCCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT	800

5	801	CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC	840
10	841	CGTGGTTCTGCCCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
15	881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
20	921	CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG	960
25	961	ATCATGGCCTCTCCAGTTGGATTGAGCGGGCCCGAGTTTA	1000
30	1001	CCTTTCCTCTCTATGGAATATGGGAAACGCCGCTCCACA	1040
35	1041	ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA	1080
40	1081	ACCTTGCTCTCCACCTTGTACAGAAGACCCCTTCAATATCG	1120
45	1121	GTATCAACAACCGAGCAACTTTCGGTCTTGACGGAACAGA	1160
50	1161	GTTCCGCTATGGAACCTCTTCTAACTTGCCATCGCGTGTT	1200
55	1201	TACAGAAAGAGCGGGAACCGTTGATTCCTTGGACGAAATCC	1240
60	1241	CACCACAGAACAAATGTGCCACCCAGGCAAGGATTCTC	1280
65	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCGGATTTC	1320
70	1321	AGCAACAGTTCGGTGAGCATCATCAGAGCTCCTATGTTCT	1360
75	1361	CATGGATTTCATCGTAGTGCTGAGTTCAACAATATCATTTCC	1400
80	1401	TTCTCTCAAATCACCCAAATCCCATTTGACCAAGTCTACT	1440
85	1441	AACCTTGGATCTTGGAACCTTCTGCTGAAAGGACCAGGCT	1480
90	1481	TCACAGGAGGTGATATTCTTAGAAGAACTTCTCCTGGCCA	1520

5 1521 GATTAGCACCCTCAGAGTTAACATCACTGCACCCTTTCT 1560  
 1561 CAAAGATATCGTGTGAGGATTGTTACGCATCTACCACTA 1600  
 10 1601 ACTTGCAATTCACACCTCCATCGACGGAAGGCCTATCAA 1640  
 1641 TCAGGGTAACTTCTCCGCAACCATGTCAAGCGGCAGCAAC 1680  
 15 1681 TTGCAATCCGGCAGCTTCAGAACCGTCGGTTTCACTACTC 1720  
 1721 CTTTCAACTTCTCTAACGGATCAAGCGTTTTTACCCCTTAG 1760  
 20 1761 CGCTCATGTGTTCAATTCTGGCAATGAAGTGTACATTGAC 1800  
 1801 CGTATTGAGTTTGTGCCTGCCGAAGTTACCTCGAGGCTG 1840  
 25 1841 AGTACAACCTTGAGAGAGCCCAGAAGGCTGTGAACGCCCT 1880  
 1881 CTTTACCTCCACCAATCAGCTTGGCTTGAAAATAACGTT 1920  
 30 1921 ACTGAATATCACATTGACCAAGTGTCCAACCTTGTCACCT 1960  
 35 1961 ACCTTAGCGATGAGTTCTGCCTCGACGAGAAGCGTGAAT 2000  
 2001 CTCCGAGAAAGTTAAACACGCCAAGCGTCTCAGCGACGAG 2040  
 40 2041 AGGAATCTCTTGCAAGACTCCAACCTCAAAGACATCAACA 2080  
 2081 GGCAGCCAGAACGTGGTTGGGGTGAAGCACCGGGATCAC 2120  
 45 2121 CATCCAAGGAGGCGACGATGTGTTCAAGGAGAATACTACGTC 2160  
 50 2161 ACCCTCTCCGGAACCTTCGACGAGTGCTACCTACCTACT 2200  
 2201 TGTACCAGAAGATCGATGAGTCCAACCTCAAAGCCTTCAC 2240  
 55 2241 CAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAAGAC 2280



2281 CTTGAAATCTACTCGATCAGGTACAATGCCAAGCACGAGA 2320  
 2321 CCGTGAATGTCCCAGGTACTGGTTCCCTCTGGCCACTTTC 2360  
 2361 TGCCCAATCTCCCATTTGGGAAGTGTGGAGAGCCTAACAGA 2400  
 2401 TGGCTCCACACCTTGAGTGAATCCTGACTTGGACTGCT 2440  
 2441 CCTGCAGGGATGGCGAGAAGTGTGCCACCATTCTCATCA 2480  
 2481 CTTCTCCTTGGACATCGATGTGGGATGTACTGACCTGAAT 2520  
 2521 GAGGACCTCGGAGTCTGGGTCTATCTTCAAGATCAAGACCC 2560  
 2561 AAGACGGACACGCAAGACTTGGCAACCTTGAGTTCTCGA 2600  
 2601 AGAGAAACCATTGGTCGGTGAAGCTCTCGCTCGTGTGAAG 2640  
 2641 AGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAACTCG 2680  
 2681 AATGGGAAACTAACATCGTTTACAAGGAGGCCAAAGAGTC 2720  
 2721 CGTGGATGCTTTTGTTCGTGAACCTCCAATATGATCAGTTG 2760  
 2761 CAAGCCGACACCAACATCGCCATGATCCAGCCGCAGACA 2800  
 2801 AACGTGTGCACAGCATTCGTGAGGCTTACTTGCCCTGAGTT 2840  
 2841 GTCCGTGATCCCTGGTGTGAACGCTGCCATCTTCGAGGAA 2880  
 2881 CTTGAGGGACGTATCTTTACCGCATTCCTTGTACGATG 2920  
 2921 CCAGAAACGTCTATCAAGAACGGTGACTTCAACAATGGCCT 2960  
 2961 CAGCTGCTGGAATGTGAAGGTCATGTGGACGTGGAGGAA 3000  
 3001 CAGAACAATCAGCGTTCCTGCTGCTGAGTGGG 3040

3041 AAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGGTTAG 3080  
 3081 AGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGATAC 3120  
 3121 GGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAACA 3160  
 3161 CCGACGAGCTTAAGTTCTCCAAGTCCGTCGAGGAAGAAAT 3200  
 3201 CTATCCCAACAACACCGTTACTTGCAACGACTACACTGTG 3240  
 3241 AATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTAACA 3280  
 3281 GAGGTTACAACGAAGCTCCTTCCGTTCCGTGCTGACTATGC 3320  
 3321 CTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGACGT 3360  
 3361 GAGAACCCTTGCGAGTTCAACAGAGGTTACAGGGACTACA 3400  
 3401 CACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGAGTA 3440  
 3441 CTTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAA 3480  
 3481 ACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTCTCT 3520  
 3521 TGATGGAGGAA 3531.

# Revendications

- Procédé de modification d'une séquence de gène de structure du type sauvage qui code une protéine insecticide de *Bacillus thuringiensis* afin d'activer l'expression de ladite protéine chez des plantes qui comprend :

- l'identification de régions à l'intérieur de ladite séquence comprenant plus de quatre nucléotides consécutifs d'adénine ou de thymine,
- la modification des régions de l'étape a) qui comportent deux ou plusieurs signaux de polyadénylation à l'intérieur d'une séquence de dix bases afin d'éliminer lesdits signaux tout en conservant une séquence de gène qui code ladite protéine, et
- la modification des régions de 15 à 30 bases entourant les régions de l'étape a) afin d'éliminer les signaux majeurs de polyadénylation de plantes, les séquences consécutives contenant plus d'un signal mineur de polyadénylation et les séquences consécutives contenant plus d'une séquence ATTTA tout en conservant une séquence de gène qui code ladite protéine.

2. Procédé de modification d'une séquence de gène de structure du type sauvage qui code une protéine insecticide de *Bacillus thuringiensis* afin d'activer l'expression de ladite protéine chez des plantes qui comprend :

- a) l'élimination des signaux de polyadénylation contenus dans ledit gène de type sauvage tout en conservant une séquence qui code ladite protéine, et  
b) l'élimination des séquences ATTTA contenues dans ledit gène de type sauvage tout en conservant une séquence qui code ladite protéine.

3. Procédé selon la revendication 2, comprenant en outre l'élimination des séquences autoccomplémentaires et le remplacement de telles séquences par de l'ADN non autoccomplémentaire comprenant des codons préférés des plantes tout en conservant une séquence de gène de structure codant ladite protéine.

4. Procédé selon les revendications 1 à 3, comprenant en outre l'utilisation des séquences préférées des plantes au cours de l'élimination des signaux de polyadénylation et des séquences ATTTA.

5. Procédé selon les revendications 1 à 3, dans lequel les signaux de polyadénylation des plantes sont choisis parmi le groupe constitué de AATAAA, AATAAT, AACCAA, ATATAA, AATCAA, ATACTA, ATAAAA, ATGAAA, AAGCAT, ATTAAT, ATACAT, AAAATA, ATTAAA, AATTAA, AATACA et CATAAA.

6. Procédé destiné à améliorer l'expression d'un gène hétérologue chez des plantes dans lequel ledit gène comprend un gène chimère modifié comprenant un promoteur qui agit dans les cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont une partie au moins est dérivée d'une protéine de *Bacillus thuringiensis*, dans lequel ledit procédé comprend la modification de ladite séquence de structure codante de sorte que ladite séquence comporte une séquence d'ADN qui diffère de la séquence d'ADN apparaissant dans la nature codant ladite protéine de *Bacillus thuringiensis* et ladite séquence de structure codante ne contient pas plus de 5 nucléotides consécutifs constitués de restes soit adénine, soit thymine.

7. Procédé d'amélioration de l'expression d'un gène hétérologue chez des plantes dans lequel ledit gène comprend un gène chimère modifié comprenant un promoteur qui agit dans des cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée d'une protéine de *Bacillus thuringiensis*, dans lequel ledit procédé comprend la modification de ladite séquence de structure codante de sorte que ladite séquence comporte une séquence d'ADN qui diffère de la séquence d'ADN qui apparaît dans la nature codant ladite protéine de *Bacillus thuringiensis* et présente les caractéristiques suivantes :

ladite séquence de structure codante comporte une région qui est complémentaire de la séquence suivante :

GGCTTGATTCCTAGCGAACTCTTCGATTCTCTGGTTGATGAGCTGTTT

1 5 10 15 20 25 30 35 40 45

ladite région dans ladite séquence codante ayant éliminé 2 séquences AACCAA et 1 séquence AATTAA.

8. Procédé selon la revendication 7, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée de *Bacillus thuringiensis* kurstaki HD-1.

9. Procédé selon la revendication 7 ou 8, dans lequel la plante est un plan de tabac.

10. Gène chimère modifié contenant un promoteur qui agit dans des cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée d'une

protéine de *Bacillus thuringiensis*, dans lequel ladite séquence de structure codante comporte une séquence d'ADN qui diffère de la séquence d'ADN apparaissant dans la nature codant ladite protéine de *Bacillus thuringiensis* et est choisie à partir de :

A. Un gène de structure qui code une protéine insecticide de *B.t.k.* HD-1 comportant la séquence :

1 ATGGCTATAGAACTGGTTACACCCCAATCGATATTTCCT 40  
 5  
 41 TGTGCTAACGCAATTTCTTTTGTAGTGAATTTGTTCCCGG 80  
 10  
 81 TGCTGGATTTGTGTTAGGACTAGTTGATATTATCTGGGGA 120  
 121 ATTTTGGTCCCTCTCAATGGGACGCATTTCTGTACAAA 160  
 15  
 161 TTGAACAGCTCATCAACCAGAGAATCGAAGAGTTCGCTAG 200  
 201 GAATCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTT 240  
 20  
 241 TATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAG 280  
 281 ATCCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCA 320  
 321 ATTCAATGACATGAACAGTGCCCTTACAACCGCTATTCCCT 360  
 30  
 361 CTTTTGCAAGTTCAAAATTATCAAGTTCCTCTCCTCTCCG 400  
 401 TGTACGTTCAAGCTGCCAACCTCCACCTCTCAGTTTGTAG 440  
 35  
 441 AGATGTTTCAGTGTTTGGACAAAGGTGGGGATTGTATGCC 480  
 481 GCGACTATCAATAGTCGTTATAATGATTTAACTAGGCTTA 520  
 521 TTGGCAACTATACAGATCATGCTGTACGCTGGTACAATAC 560  
 45  
 561 GGGATTAGAGCGTGTATGGGGACCGGATTCTAGAGATTGG 600  
 601 ATCAGGTACAACCAGTTCAGAAGAGAGCTTACACTAACTG 640  
 50  
 641 TATTAGATATCGTTTCTCTATTTCGGAAGCTATGATAGTAG 680  
 681 AACGTATCCAATTGGAACAGTTTCGCAATTAACAAGAGAA 720  
 55

5 721 ATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTT 760  
 761 TTCGAGGCTCGGGCTCAGGGCATAGAAGGAAGTATTAGGAG 800  
 10 801 TCCACATTTGATGGATATACTTAATAGTATAACCATCTAT 840  
 841 ACGGATGCTCATAGAGGAGAATACTACTGGTCCGGTCACC 880  
 15 881 AGATCATGGCTTCTCCTGTAGGGTTTTTCGGGGCCAGAATT 920  
 20 921 CACTTTTCCGCTATATGGAACATGCGGAAATGCAGCTCCA 960  
 961 CAACAACGTATTGTTGCTCAACTAGGTCAGGGCGGTGTATA 1000  
 25 1001 GAACATTATCGTCCACCTTATATAGAAGACCTTTTAACAT 1040  
 1041 CGGGATCAACAACCAACAACCTATCTGTTCTTGACGGGACA 1080  
 30 1081 GAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTG 1120

1121 TATACAGAAAAAGCGGAACGGTAGATTGCGCTGGATGAAAT 1160  
 1161 ACCGCCACAGAATAACAACGTGCCACCTAGGCCAAGGATTT 1200  
 1201 AGTCATCGATTAGCCATGTTTCAATGTTTCGTTTCAGGCT 1240  
 1241 TTAGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTT 1280  
 1281 CTCTTGGATACATCGTAGTGCTGAGTTCAACAACATCATC 1320  
 1321 CCTTCATCACAAATCACCCAAATCCCCTCACCAGTCTA 1360  
 1361 CTAATCTTGGCTCTGGAACTTCTGTCTGTTAAAGGACCAGG 1400  
 1401 ATTTACAGGAGGAGATATTCTTCGAAGAATTCACCTGGC 1440  
 1441 CAGATTTCAACCTTAAGAGTAAATATTACTGCACCATTAT 1480  
 1481 CACAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCAC 1520  
 1521 AAACCTTCAGTTCCACACATCAATTGACGGAAGACCTATT 1560  
 1561 AATCAGGGGAATTTTTCAGCAACTATGAGTAGTGGGAGTA 1600  
 1601 ATTTACAGTCCGGAAGCTTTAGGACTGTAGGTTTACTAC 1640  
 1641 TCCGTTTAACTTTTCAAATGGATCAAGTGTATTTACGTTA 1680  
 1681 AGTGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAG 1720  
 1721 ATCGAATTGAATTGTTCGGCA 1743.

B. Un gène de structure qui code une protéine insecticide de *B.t.k.* HD-73 comportant la séquence :

1 ATGGCCATTGAAACCGGTTACACTCCCATCGACATCTCCT 40  
 5  
 41 TGTCCCTTGACACAGTTTCTGCTCAGCGAGTTGCTGCCAGG 80  
 10  
 81 TGCTGGGTTTCGTTCTCGGACTAGTTGACATCATCTGGGGT 120  
 121 ATCTTTGGTCCATCTCAATGGGATGCATTCCCTGGTGCAAA 160  
 15  
 161 TTGAGCAGTTGATCAACCAGAGGATCGAAGAGTTGCCCAG 200  
 201 GAACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTC 240  
 20  
 241 TACCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCG 280  
 25  
 281 ATCCTACTAATCCAGCTCTCCGCGAGGAAATGCGTATTCA 320  
 321 ATTCAACGACATGAACAGCGCCTTGACCACAGCTATCCCA 360  
 30  
 361 TTGTTGCGAGTCCAGAACTACCAAGTTCTCTCTTGTCCG 400  
 35  
 401 TGTACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCG 440



441 AGACGTTAGCGTGTTTGGGCAAAGGTGGGAATTCGATGCT 480  
 481 GCAACCATCAATAGCCGTTACAACGACCTTACTAGGCTGA 520  
 521 TTGGAAACTACACCGACCACGCTGTTGCTTGGTACAACAC 560  
 561 TGGCTTGGAGCGTGTCTGGGGTCCCTGATCTAGAGATTGG 600  
 601 ATTAGATACAACCAAGTTCAGGAGAGAATTGACCCTCACAG 640  
 641 TTTTGGACATTGTGTCTCTCTCCCGAACTATGACTCCAG 680  
 681 AACCTACCCTATCCGTACAGTGTCCCAACTTACCAGAGAA 720  
 721 ATCTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCT 760  
 751 TCCGTGGTTCTGCCCCAAGGTATCGAAGGCTCCATCAGGAG 800  
 801 CCCACACTTGTATGGACATCTTGAACAGCATAACTATCTAC 840  
 841 ACCGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACC 880  
 881 AGATCATGGCCTCTCCAGTTGGATTGACGGGGCCCGAGTT 920  
 921 TACCTTTCCTCTCTATGGAAGTATGGGAAACGCCGCTCCA 960  
 961 CAACAACGTATCGTTGCTCAACTAGGTCAGGGGTGTCTACA 1000  
 1001 GAACCTTGTCTTCCACCTTGTACAGAGACCCCTTCAATAT 1040

5 1041 CGGTATCAACAACCAGCAACTTTCCGGTTCTTGACGGAACA 1080  
 1081 GAGTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTG 1120  
 10 1121 TTACAGAAAGAGCGGAACCGTTGATTCCCTGGACGAAAT 1160  
 1161 CCCACCACAGACAACAATGTGCCACCCAGGCAAGGATT 1200  
 15 1201 TCCACAGGTTGAGCCACGTGTCCATGTTCGGTCCGGAT 1240  
 1241 TCAGCAACAGTTCGGTGAGCATCATCAGAGCTCCTATGTT 1280  
 20 1281 CTCTTGATACACCGTAGTGCTGAGTTCAACAACATCATC 1320  
 1321 GCATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAA 1360  
 25 1361 ACTTTCTCTTCAACGGTTCTGTCAATTCCAGGACCAGGATT 1400  
 30 1401 CACTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAAT 1440  
 1441 AACATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACT 1480  
 35 1481 TCCCATCCACATCTACCAGATATAGASTTCGTGTSAGGTA 1520  
 40 1521 TGCTTCTGTGACCCCTATTCACTCAACGTTAATTGGGGT 1560  
 1561 AATTGATCCATCTTCTCCAATACAGTTCAGCTACAGCTA 1600  
 45 1601 CCTCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTT 1640

50

55

1641 TGAAAGTGCCAATGCTTTTACATCTTCACTCGGTAACATC 1680  
 1681 GTGGGTGTTAGAAACTTTAGTGGGACTGCAGGAGTGATTA 1720  
 1721 TCGACAGATTCGAGTTCATTCCAGTTACTGCAACACTCGA 1760  
 1761 GGCTGAG 1767.

C. Un gène de structure codant une protéine insecticide de *B.t.k.* HD-1 comportant la séquence :

1 ATGGACAACACCCAAACATCAACGAATGCATTCCATACA 40  
 41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGSAGA 80  
 81 ACGCATTGAAACCGGTTTACACTCCCATCGACATCTCCTTG 120  
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160  
 161 CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200  
 201 CTTTGGTCCATCTCAATGGGATGCATTCCCTGGTGCAAATT 240  
 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTGCCCAGGA 280  
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320  
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360

361 CCTACTAACCCAGCTCTCCGCCAGGAAATGCGTATTCAAT 400  
 401 TCAACGACATGAACAGCGCCTTGACCAAGCTATCCCATTT 440  
 441 GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGGTG 480  
 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520  
 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGATTGATGCTGC 560  
 561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600  
 601 GGAAACTACACCGACCGCTGTTGTTGGTACAACTG 640  
 641 GCTTGGAGCGTGCTGCGGTCTGATTCTAGAGATTGGAT 680  
 681 TAGATACAACCAGTTCAGGAGAGAATTGACCCCTCACAGTT 720  
 721 TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA 760  
 761 CCTACCCATATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800  
 801 CTATACTAACCCAGTTCCTTGAGAACTTCGACGGTAGCTTC 840  
 841 CGTGGTTCTGCCCAAGGTATCGAAGGCTCCATCAGGAGCC 880  
 881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920  
 921 CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG 960

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

961 ATCATGGCCTCTCCAGTTGGATTCAACGGGCCCCGAGTTTA 1000  
 1001 CCTTTCCTCTCTATGGAACATATGGGAACGCCGCTCCACA 1040  
 1041 ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA 1080  
 1081 ACCTTGTCTTCCACCTTGTACAGAAGACCCTTCAATATCG 1120  
 1121 GTATCAACAACCAGCAACTTTCCGTTCTTGACGGAACAGA 1160  
 1161 GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT 1200  
 1201 TACAGAAAGAGCGGAACCGTTGATTCTTGACGAAATCC 1240  
 1241 CACCACAGAAACAATGTGCCACCCAGGCAAGGATTCTC 1280  
 1281 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTCT 1320  
 1321 AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT 1360  
 1361 CATGGATTCAATCGTAGTGCTGAGTTCAACAATATCATTC 1400  
 1401 TTCCTCTCAAATCACCCAAATCCATTGACCAAGTCTACT 1440  
 1441 AACCTTGGATCTGGAACCTCTGTGCTGAAAGGACCAGGCT 1480  
 1481 TCACAGGAGGTGATATTCTTAGAAGAAGTCTCTCOTGGCCA 1520  
 1521 GATTAGCACCCCTCAGAGTTAAACATCACTGCACCACCTTCT 1560

1561 CAAAGATATCGTGTCAAGGATTCGTTACGCATCTACCACTA 1600  
 1601 ACTTGCAATTCCACACCTCCATCGACGGAAGGCCTATCAA 1640  
 1641 TCAGSGTAACCTTCTCCGCAACCATGTCAAGCGGCAGCAAC 1680  
 1681 TTGCAATCCGGCAGCTTCAGAACCCTCGGTTTCACTACTC 1720  
 1721 CTTTCAACTTCTCTAACGGATCAAGCGTTTTCACCCCTTAG 1760  
 1761 CGCTCATGTGTTCAATTCTGGCAATGAAGTGATCATTTGAC 1800  
 1801 CGTATTGAGTTTGTGCCTGCCGAAGTTACCTTCGAGGCTG 1840  
 1841 AGTAC 1845.

D. Un gène de structure codant une protéine insecticide dérivée de *B.t.k.* HD-73 comportant la séquence :

1 ATGGACAACAACCCAAACATCAACGAATGCAATCCATACA 40  
 41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGCTGGAGA 80  
 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120  
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160  
 161 CTGGGTTCTGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200

201 CTTTGGTCCATCTCAATGGGATGCAATTCCTGGTGCAAATT 240  
 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCCGCCAGGA 280  
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320  
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360  
 361 CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT 400  
 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCAAT 440  
 441 GTTCGCAGTCCAGAACTACCAAGTTCTCTCTCTGTCCGTG 480  
 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520  
 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTGATGCTGC 560  
 561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600  
 601 GGAAACTACACCGACCACGCTGTTTCGTTGGTACAACACTG 640  
 641 GCTTGGAGCGTGCTGCGGTCTGATTCTAGAGATTGGAT 680  
 681 TAGATACAACCAGTTTCAGGAGAGAATTGACCCCTCACAGTT 720  
 721 TTGGACATTGTGTCTCTCTTCCCCGAACATGACTCCAGAA 760  
 761 CCTACCCATCCGTTACAGTGTCCCAACTTACCAGAGAAAT 800

5	801	CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC	840
	841	CGTGGTTCTGCCCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
10	881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
15	921	CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG	960
20	961	ATCATGGCCTCTCCAGTTGGATTACGCGGCCCCGAGTTTA	1000
	1001	CCTTTCTCTCTATGGAACATGCGGAACGCCGCTCCACA	1040
25	1041	ACAACGTATCGTTGCTCAACTAGGTGAGGGTGTCTACAGA	1080
30	1081	ACCTTGCTCTTCCACCTTGACAGAAGACCCCTTCAATATCG	1120
	1121	GTATCAACAACCAGCAACTTCCGTTCTTGACGGAACAGA	1160
35	1161	GTTGCGCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT	1200
	1201	TACAGAAAGAGCGGAACCGTTGATTCCITGGACGAAATCC	1240
40	1241	CACCACAGAACACAATGTGCCACCAGGCAAGGATTCTC	1280
	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC	1320
45	1321	AGCAACAGTTCGCTGAGCATCATCAGAGCTCCTATGTTCT	1360
50	1361	CTTGGATACACCGTAGTGCTGAGTTCAACAACATCATCGC	1400

55



1401 ATCCGATAGTATTACTCAAATCCCTGCACTGAAGGGAAC 1440  
 1441 TTTCTCTTCAACGGTTCTGTTCATTTTCAGGACCAGGATTCA 1480  
 1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520  
 1521 CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACCTC 1560  
 1561 CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG 1600  
 1601 CTTCTGTGACCCCTATTCACTCAACGTTAATTGGGGTAA 1640  
 1641 TTCATCCATCTTCTCCAATACAGTTCAGGTACAGCTACC 1680  
 1681 TCCTTGGATAATCTCCAATCCAGCGATTTCCGGTTACTTTG 1720  
 1721 AAAGTGCCAAATGCTTTTACATCTTCACTCGGTAACATCGT 1760  
 1761 GGGTGTTAGAAACTTTAGTGGGACTGCAAGGAGTGATTATC 1800  
 1801 GACAGATTCGAGTTCATTCAGTTACTGCAACACTCGAGG 1840  
 1841 CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTAAATGCG 1880  
 1881 CTGTTTACGTCTACAAACCAGCTTGGACTCAAGACAAATG 1920  
 1921 G 1921,

E. Un gène de structure codant la protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

1 ATGGACAACACCCCAACATCAACGAATGCATTCCATACA 40  
 41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA 80  
 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120  
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160  
 161 CTGGGTTGCTTCTCGGACTAGTTGACATCATCTGGGGTAT 200  
 201 CTTTGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATT 240  
 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTTCGCCAGGA 280  
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320  
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360  
 361 CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT 400  
 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCAT 440  
 441 GTTCGCAGTCCAGAACTACCAAGTTCCCTCTCTTGTCCGTG 480  
 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520  
 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTGATGCTGC 560  
 561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600

601 GGAAACTACACCGACCACGCTGTCGTTGGTACAACACTG 640  
 641 GCTTGGAGCGTGTCTGGGGTCCTGATTCTAGAGATTGGAT 680  
 681 TAGATACAACCCAGTTCAGGAGAGAATTGACCCCTCACAGTT 720  
 721 TTGGACATTGTGTCTCTCTTCCCGAAGTATGACTCCAGAA 760  
 761 CCTACCCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800  
 801 CTATACTAACCAGTTCCTTGAGAACTTCGACGGTAGCTTC 840  
 841 CGTGGTTCTGCCCCAAGGTATCGAAGGCTCCATCAGGAGCC 880  
 881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920  
 921 CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG 960  
 961 ATCATGGCCTCTCCAGTTGSAITTCAGCGGGCCCCGAGTTTA 1000  
 1001 CCTTTCCTCTCTATGGAAGTATGGGAAACGCCGCTCCACA 1040  
 1041 ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA 1080  
 1081 ACCTTGTCTTCCACCTTGTACAGAAGACCCCTTCAATATCG 1120  
 1121 GTATCAACAACAGCAACTTTCGGTTCCTTGACGGAACAGA 1160  
 1161 GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT 1200

1201 TACAGAAAGAGCGGGAACCGTTGATTCCCTTGGACGAAATCC 1240  
 1241 CACCACAGAAACAACATGTGCCACCCAGGCAAGGATTCTC 1280  
 1281 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTCC 1320  
 1321 AGCAACAGTTCGCGTGAGCATCATCAGAGCTCCTATGTTCT 1360  
 1361 CTTGGATACACCGTAGTGCTGAGTTCAACAACATCATCGC 1400  
 1401 ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGAAAC 1440  
 1441 TTTCTCTTCAACGGTCTCTGTCAATTCAGGACCAGGATTCA 1480  
 1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520  
 1521 CATTGAGAATAGAGGGTATATTGAAGTTCCAATTCAC TTC 1560  
 1561 CCATCCACATCTACCCAGATATAGAGTTCTGTGTGAGGTATG 1600  
 1601 CTTCTGTGACCCCTATTCAACCTCAACGTTAATTGGGGTAA 1640  
 1641 TTCATCCATCTTCTCCAATACAGTTCAGCTACAGCTACC 1680  
 1681 TCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTTTG 1720  
 1721 AAAGTGCCAATGCTTTTACATCTTCACTCGGTAAACATCGT 1760  
 1761 GGGTGTGTAGAACTTTAGTGGGACTGCAGGAGTGATTATC 1800

5 1801 GACAGATTTCGAGTTCATTCCAGTTACTGCAACACTCGAGG 1540  
 1841 CTGAATATAATCTGGAAGAGCGCAGAAGCGGTGAATGC 1880  
 10 1881 GCTGTTTACGTCTACAAACCAGCTCGGCCCTCAAGACCAAT 1920  
 1921 GTGACGGATTATCATATTGATCAAGTGTCCAACCTGGTGA 1960  
 15  
 1961 CCTACCTCAGCGATGAGTTCTGTCTGGATGAAAAGCGAGA 2000  
 20 2001 ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGA 2040  
 2041 GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTAA 2080  
 2081 ATAGGCAACCAGAACGTGGGTGGGGCGGAAGTACAGGGAT 2120  
 30 2121 TACCATCCAGGGAGGTGACGACGTGTTCAAGGAGAACTAC 2160  
 2161 GTCACACTATCAGGTACCTTTTGATGAGTGCTATCCAACAT 2200  
 35 2201 ACCTCTACCAGAAGATCGACGAGTCCAAGTTGAAAAGCCTT 2240  
 2241 TACCCGTTATCAATTAAAGAGGTATATCGAAGATAGTCAA 2280  
 40 2281 GACCTCGAGATCTACCTCATCCGCTACAATGCAAAACATG 2320  
 2321 AAACAGTAAATGTGCCAGGTACGGGTTCCCTTATGGCCGCT 2360  
 45 2361 TTCAGCCCAAAGTCCCAATCGGAAAAGTGTGGAGAGCCGAAT 2400  
 50  
 55

5 2401 CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT 2440  
 2441 GTTCGTGTAGGGATGGAGAAAAAGTGTGCCCATCATTTCGCA 2480  
 10 2481 TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA 2520  
 2521 AATGAGGACCTAGGTGTATGGGTGATCTTAAAGATTAA 2560  
 15 2561 CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT 2600  
 2601 CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG 2640  
 2641 AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAGAAGT 2680  
 25 2681 TGGAATGGGAGACCAACATCGTCTACAAGAGGCAAAAGA 2720  
 2721 ATCTGTAGATGCTTTATTGTAACTCTCAATATGATCAA 2760  
 2761 TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG 2800  
 30 2801 ATAAACGTGTTTCATAGCATTCGAGAAGCTTATCTGCCTGA 2840  
 2841 GCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTGA 2880  
 35 2881 GAATTAGAAGGGCGTATTTTCACTGCAATCTCCCTCTACG 2920  
 2921 ATGCCAGAAACGTCATCAAGAACGCTGACTTCAACAATGG 2960  
 40 2961 CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA 3000  
 45  
 50  
 55

3001 GAACAAACAACCAACGTTCCGGTCCTTGTGTGTTCCGGAAT 3040  
 3041 GGGGARGCAGAAGTGTCCACAGAAGTTCGTGTCTGTCCGGG 3080  
 3081 TCGTGGCTATATCCTTCGTGTCCACGCGTACAAGGAGGGA 3120  
 3121 TATGGAGAAGGTTGCGTAACCATTCATGAGATCGAGAACA 3160  
 3161 ATACAGACGAACTGAAGTTTAGCAACTCCGTAGAAGAGGA 3200  
 3201 AATCTATCCAAATAACACGGTAACGTGTAAATGATTATACT 3240  
 3241 GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA 3280  
 3281 ATCGAGGATATAACGAAGCTCCTTCCGTACCAGCTGATTA 3320  
 3321 TCGCTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA 3360  
 3361 AGAGAGAATCCTTGTGAATTAAACAGAGGGTATAGGGAAT 3400  
 3401 ACACGCCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA 3440  
 3441 ATACTTCCCAGAAACCGATTAAGGTATGGATTGAGATTGGA 3480  
 3481 GAAACGGAAGGAACATTATCGTGGACAGCGTGAATTAC 3520  
 3521 TCCTTATGGAGGAA 3534.

F. Un gène de structure codant une protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40  
 41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA 80  
 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120  
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160  
 161 CTGGGTTTCGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200  
 201 CTTTGGTCCATCTCAATGGGATGCATTCTCTGGTGCAAATT 240  
 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA 280  
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320  
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360  
 361 CCTACTAACCAGCTCTCCGCGAGGAAATGCGTATTCAAT 400  
 401 TCAACGACATGAACAGCGCCCTTGACCACAGCTATCCCATT 440  
 441 GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG 480  
 481 TACGTTCAAGCAGCTAATCTTCACTCAGCGTGCTTCGAG 520  
 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTGATGCTGC 560



561 AACATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 500  
 5 GGAAGTACACCGACCGCTGTTCTGTTGGTACAACACTG 640  
 641 GCTTGGAGCGTGTCTGGGGTCTCTGATTCTAGAGATTGGAT 680  
 10 TAGATACAACCAAGTTTCAGGAGAGAATTGACCCCTCACAGTT 720  
 721 TTGGACATTGTGTCTCTCTTCCCGAAGTATGACTCCAGAA 760  
 15 CCTACCCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800  
 20 CTATACTAAGCCAGTTCTTTGAGAAGTTGACGCGTAGCTTC 840  
 841 CGTGGTTCGCCCCAAGGTATCGAAGGCTCCATCAGGAGCC 880  
 25 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920  
 921 CGATGCTCAGAGGAGAGTATTACTGGTCTGGACACCCAG 960  
 30 ATCATGGCCCTCTCCAGTTGGATTTCAGCGGGCCCGAGTTTA 1000  
 35 CCTTTCCTCTCTATGGAAGTATGGGAAACGCGGCTCCACA 1040  
 1041 ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA 1080  
 40 ACCTTGCTCTCCACCTTGTACAGAAGACCCCTTCAATATCG 1120  
 1121 GTATCAACAACCAAGCACTTTCCGTTCTTTGACGGAACAGA 1160  
 45 GTTCGCCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT 1200  
 50 TACAGAAAGAGCGGAACCGTTGATTCTTTGGACGAAATCC 1240  
 1241 CACCACAGAACAACTATGTGCCACCCAGGCAAGGATTCTC 1280  
 55 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC 1320

1321 AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT 1360  
 1361 CTGGGATACACCGTAGTGCTGAGTTCAACAACATCATCGC 1400  
 1401 ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC 1440  
 1441 TTTCTCTTCAACGGTTCTGTCAATTCAGGACCAGGATTCA 1480  
 1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520  
 1521 CATTGAGATAGAGGGTATATTGAAGTTCCAATTCAC TTC 1560  
 1561 CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG 1600  
 1601 CTTCTGTGACCCCTATTACCTCAACGTTAATTGGGGTAA 1640  
 1641 TTCAATCCATCTTCTCCAATACAGTTCCAGCTACAGCTACC 1680  
 1681 TCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTTTG 1720  
 1721 AAAGTGCCAATGCTTTTACATCTTCACTCGGTAACATCGT 1760  
 1761 GGGTGTTAGAAACTTTAGTGGGACTGCAGGAGTGATTATC 1800  
 1801 GACAGATTGAGTTCAATCCAGTTACTGCAACACTCGAGG 1840  
 1841 CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC 1880  
 1881 GCTGTTTACGTCTACAAACCACTAGGGCTAAAAACAAAT 1920  
 1921 GTAACGGATTATCATATTGATCAAGTGTCCAATTTAGTTA 1960  
 1961 CGTATTATCGGATGAATTTTGTCTGGATGAAAAGCGAGA 2000  
 2001 ATTGTCGAGAAAGTCAAAACATGCGAAGGACTCAGTGAT 2040  
 2041 GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTA 2080

2081 ATAGGCAACCGAACCTGGGTGGGGCGGAAGTACAGGGAT 2120  
 2121 TACCATCCAAGGAGGGGATGACGTATTTAAAGAAAATTAC 2160  
 2161 GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT 2200  
 2201 ATTTGTATCAAAAAATCGATGAATCAAAATTAAAGCCTT 2240  
 2241 TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA 2280  
 2281 GACTTAGAAATCTATTTAATTCGCTACAATGCAAAACATG 2320  
 2321 AAACAGTAAATGTGCCAGGTACGGGTTCCCTTATGGCCGCT 2360  
 2361 TTCAGCCCAAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT 2400  
 2401 CGATCGCGGCCACACCTTGAATGGAATCCTGACTTAGATT 2440  
 2441 GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATCATTCGCA 2480  
 2481 TCATTTCCTCTTAGACATTGATGTAGGATGTACAGACTTA 2520  
 2521 AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAAAG 2560  
 2561 CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT 2600  
 2601 CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG 2640  
 2641 AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAAAAAT 2680  
 2681 TGGAAATGGGAAACAAATATCGTTTATAAAGAGGCAAAAGA 2720  
 2721 ATCTGTAGATGCTTTATTGTAAACTCTCAATATGATCAA 2760  
 2761 TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG 2800  
 2801 ATAAACGTGTTTCATAGCATTGAGAGAAGCTTATCTGCCTGA 2840

2841 GCGTCTGTGATTCGCGGGTGTCAATGCGGCTATTTTGA 2880  
 5  
 2881 GAATTAGAAGGGCGTATTTTCACTGCATTCTCCCTATATG 2920  
 10  
 2921 ATGCGAGAAATGTCATTAATAATGGTGATTTTAAATAATGG 2960  
 15  
 2961 CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA 3000  
 20  
 3001 GAACAAAACAACCAACGTTTCGGTCCCTTGTGTTCCGGAAT 3040  
 25  
 3041 GGGAAAGCAGAGTGTCAAGAAGTTCGTGCTGTCCGGG 3080  
 3081 TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGGA 3120  
 3121 TATGGAGAAGGTTGCGTAACCATTCATGAGATCGAGAACA 3160  
 3161 ATACAGACGAACTGAAGTTTAGCAACTGCGTAGAAGAGGA 3200  
 3201 AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT 3240  
 3241 GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA 3280  
 3281 ATCGAGGATATAACGAAGCTCCTTCCGTACCAGCTGATT 3320  
 3321 TCGCTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA 3360  
 3361 AGAGAGAATCCTTGTGAATTTAAGAGGGTATAGGGATT 3400  
 3401 ACACGCCACTACCAGTTGGTTATGTGACAAAAGAAATTAGA 3440  
 3441 ATACTTCCAGAAACCGATAAGGTATGGATTGAGATTGGA 3480  
 3481 GAAACGGAAGGAACATTTATCCTGGACAGCGTGAATTAC 3520  
 3521 TCCTTATGGAGGAA 3534,

G. Un gène de structure codant une protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

```

5      1  ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA      40
      41  ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA      80
10
      81  ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG      120
      121  TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG      160
15
      161  CTGGGTTCTGTTCTCGGACTAGTTGACATCATCTGGGGTAT      200
      201  CTTTGGTCCATCTCAATGGGATGCATTCTCTGGTGCAAATT      240
20
      241  GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCCGCCAGGA      280
      281  ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA      320
25
      321  CCAAACTCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT      360
      361  CCTACTAAGCCAGCTCTCCGCGAGGAAATGCGTATTCAAT      400
30
      401  TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT      440
      441  GTTCGCAGTCCAGAAGTACCAAGTTCCCTCTCTTGTCCGTG      480
35
      481  TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG      520
      521  ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTGCGATGCTGC      560
40
      561  AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT      600
      601  GGAAACTACACCGACACGCTGTTCGTTGGTACAACACTG      640
45
      641  GCTTGGAGCGTCTCTGGGGTCCCTGATTCTAGAGATTGGAT      680
50
55

```

5

681 TAGATACAACCAGTTCCAGGAGAGAATTGACCCCTCACAGTT 720

721 TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA 760

10

761 CCTACCCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800

801 CTATACTAACCAGTTCTTGAGAACTTCGACGGTAGCTTC 840

15

841 CGTGGTTCGTGCCAAGGTATCGAAGGCTCCATCAGGAGCC 880

881 CACACTTGATGGACATCTTGAACAGCATACTATCTACAC 920

20

921 CGATGCTCAGAGAGGAGAGTATTACTGGTCTGGACACCAG 960

961 ATCATGGCCCTCTCCAGTTGGATTACGCGGGCCCGAGTTA 1000

25

1001 CCTTTCCTCTCTATGGAAGTATGGGAAACGCCGCTCCACA 1040

30

1041 ACAACGTATCGTTGCTCAACTAGGTGAGGGTGTCTACAGA 1080

1081 ACCTTGCTCTCCACCTTGATACAGAAGACCCCTTCAATATCG 1120

35

1121 GTATCAACAACAGCAACTTTCCGTTCTTGACGGAAACAGA 1160

1161 GTTCGCCATGGAACCTCTCTTAACITGCCATCCGCTGTT 1200

40

1201 TACAGAAAGAGCGGAACCGTTGATTCCCTGGACGAAATCC 1240

1241 CACCACAGAACAAATGTGCCACCCAGGCAAGGATTCTC 1280

45

1281 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC 1320

50

1321 AGCAACAGTTCGCTGAGCATCATCAGAGCTCCTATGTTCT 1360

1361 CTTGGATACACCGTAGTGCTGAGTTCAACAACATCATCGC 1400

55

1401 ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC 1440

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

```

1441 TTCTCTTCAACGGTTCGTGTCATTTCAGGACCAGGATTCA 1480
1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520
1521 CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACCTTC 1560
1561 CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG 1600
1601 CTTCTGTGACCCCTATTACCTCAACGTTAATTGGGGTAA 1640
1641 TTCATCCATCTTCTCCAATACAGTTCACGCTACAGCTACC 1680
1681 TCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTTTG 1720
1721 AAAGTGCCAAATGCTTTTACATCTTCACTCGGTAAACGCT 1760
1761 GGGTGTTAGAAACTTTAGTGGGACTGCAGGAGTGATTATC 1800
1801 GACAGATTTCGAGTTCAATCCAGTTACTGCAACACTCGAGG 1840
1841 CTGASTACAACTTGGAGAGAGCCCGAAGGCTGTGAACGC 1880
1881 CCTCTTTACCTCCACCAATCAGCTTGGCTTGAAAACCTAAC 1920
1921 GTTACTGACTATCACATTGACCAAGTGTCCAACTTGGTCA 1960
1961 CCTACCTTAGCGATGAGTTCTGCCTCGAGAGAGCGGTGA 2000
2001 ACTCTCCGAGAAAGTTAAACACGCCAAGCGTCTCAGCGAC 2040
2041 GAGAGGAATCTCTTGCAAGACTCCAACTTCAAGACATCA 2080
2081 ACAGGCAGCCAGAACGTTGGTGGGGTGGAGGCACCGGGAT 2120
2121 CACCATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTAC 2160
2161 GTCACCTCTCCGGAACCTTTCGACGAGTGCTACCCCTACCT 2200

```

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

2201	ACTTGTACCAGAAGATCGATGAGTCCAAACTCAAAGCCTT	2240
2241	CACCAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAA	2280
2281	GACCTTGRAATCTACTCGATCAGGTACAATGCCAAGCAGC	2320
2321	AGACCGTGAATGTCCCAGGTACTGGTTCCCTCTGGSCACT	2360
2361	TTCTGCCCAATCTCCCATGGGAAGTGTGGAGAGCCTAAC	2400
2401	AGATGGCGCTCCACACCTTGAGTGGAACTCCTGACTTGGACT	2440
2441	GCTCCTGCAGGGATGGCGAGAAGTGTGCCCAACCATCTCA	2480
2481	TCACTTCTCCTTGGACATCGATGTGGGATGTACTGACCTG	2520
2521	AATGAGGACCTCGGAGTCTGGGTCTCTTCAAGATCAAGA	2560
2561	CCCAAGACGGACACGCAAGACTTGGCAACCTTGAGTTTCT	2600
2601	CGAAGAGAAACCATTTGGTCGGTGAAGCTCTCGCTCGTGTG	2640
2641	AAGAGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAAAC	2680
2681	TCGAATGGGAAACTAACATCGTTTACAAGGAGGCCAAGA	2720
2721	GTCCGTGGATGCTTTGTTCTGCTGAACCTCCCAATATGATCAG	2760
2761	TTGCAAGCCGACACCAACATCGCCATGATCCACGCCGCAG	2800
2801	ACAAACGTGTGCACAGCATTCTGTGAGGCTTACTTGCCTGA	2840
2841	GTTGTCCGTGATCCCTGGTGTGAACGCTGCCATCTTCGAG	2880
2881	GAACTTGAGGGACGTATCTTTACCGCATCTCTCCTTGTACG	2920
2921	ATGCCAGAAACGTCTATCAAGAACGGTGAATTCACCAATGG	2960



```

      .
      .
      .
2961 CCTCAGCTGCTGGAATGTGAAAGGT CATGTGGACGTGGAG 3000
      .
      .
      .
3001 GAACAGAACCAATCAGCGTTCCGTCCTGGTTGTGCCTGAGT 3040
      .
      .
      .
10 3041 GGGAAAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGG 3080
      .
      .
      .
3081 TAGAGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGA 3120
      .
      .
      .
15 3121 TACGGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACA 3160
      .
      .
      .
20 3161 ACACCGACGAGCTTAAGTTCTCCAACTGCGTCGAGGAAGA 3200
      .
      .
      .
3201 AATCTATCCCAACAACACCGTTACTTGCAACGACTACACT 3240
      .
      .
      .
25 3241 GTGAATCAGGAAGAGTACGGAGGTGCCCTACACTAGCCGTA 3280
      .
      .
      .
3281 ACAGAGGTTACAACGAAGCTCCTTCCGTTCTGCTGACTA 3320
      .
      .
      .
30 3321 TGCCTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGA 3360
      .
      .
      .
35 3361 CGTGAGAACCCTTGCGAGTTCAACAGAGGTTACAGGGA CT 3400
      .
      .
      .
3401 ACACACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGA 3440
      .
      .
      .
40 3441 GTACTTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGT 3480
      .
      .
      .
45 3481 GAAACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTC 3520
      .
      .
      .
3521 TCTTGATGGAGGAA 3534,

```

H. Un gène de structure qui code une protéine insecticide de *B.t.t.* Comportant La séquence :

1 ATGACTGCAGACAACAACACCGAAGCCCTCGACAGTTCTA 40  
 5  
 41 CCACTAAGGATGTTATCCAGAAGGGTATCTCCGTTGTGGG 80  
 10  
 81 AGACCTCTTGGGCGTGGTTGGATTCCCTTCGGTGGAGCC 120  
 121 CTCGTGAGCTTCTATACAAACTTCTCAACACCAATTGGC 160  
 15  
 161 CAAGCGAGGACCCCTTGGAAAGCATTTCATGGAGCAAGTTGA 200  
 201 AGCTCTTATGGATCAGAAGATTGCAGATTATGCCAAGAAC 240  
 20  
 241 AAGGCTTTGGCAGAAGCTCCAGGGCCCTCAGAACAATGTGG 280  
 25  
 281 AGGACTACGTGAGTGCATTGTCCAGCTGGCAGAAGAACCC 320  
 321 TGTTAGCTCCAGAAATCCTCACAGCCCAAGGTAGGATCAGA 360  
 30  
 361 GAGTTGTTCTCTCAAGCCGAATCCCACTTCAGAAATTCCA 400  
 35  
 401 TGCCTAGCTTTGCTATCTCCGGTTACGAGGTTCTTTTCCT 440  
 441 CACTACCTATGCTCAAGCTGCCAACACCCACTTGTTCCTC 480  
 40  
 481 CTTAAGGACGCTCAAATCTATGGAGAAGAGTGGGGATACG 520  
 521 AGAAAGAGGACATTGCTGAGTTCTACAAGCGTCAACTTAA 560  
 45  
 561 GCTCACCCAGAGGTACACTGACCAATTGCGTGAAATGGTAT 600  
 50  
 601 AACGTTGGTCTCGATAAGCTCAGAGGCTCTTCTCAGAGT 640  
 641 CTTGGGTGAACCTTCAACAGATACAGGAGAGAGATGACCTT 680  
 55

5	681	GACTGTGCTCGATCTTATCGCACTCTTTCCCTTGACGAT	720
10	721	GTGAGACTCTACCCAAAGGAAGTGAAAAGTGAGCTTACCA	760
15	761	GAGACGTGCTCACTGACCCCTATTGTGCGGAGTCAACAACCT	800
20	801	TAGGGGTTATGGAACTACCTTCAGCAATATCGAAAACCTAC	840
25	841	ATTAGGAAACCCACATCTCTTCGACTATCTTCACAGAATTC	880
30	881	AATTCCACACAGGTTTCAACCAGGATACTATGGTAACGA	920
35	921	CTCCTTCAACTATTGGTCCGGTAACTATGTTTCCACCAGA	960
40	961	CCAGCATTTGGATCTAATGACATCATCACATCTCCCTTCT	1000
45	1001	ATGGTAACAAAGTCCAGTGAACCTGTGCAGACCTTGAGTT	1040
50	1041	CAACGGCGAGAAAGTCTATAGAGCCGTGCGAAACACCAAT	1080
55	1081	CTCGCTGTGTGGCCATCCGCAGTTTACTCAGGCGTCACAA	1120
60	1121	AGGTGGAGTTTAGTCAGTATAACGATCAGACCGATGAGGC	1160
65	1161	CAGCACCAGACTTACGACTCCAAAACGTAAAGTTGGCGCA	1200
70	1201	GTCTCTTGGGATTCTATCGACCAATTGCCCTCCGAAACCA	1240
75	1241	CAGACGAACCAATTGGAGAAGGGCTACAGCCACCAACTTAA	1280
80	1281	CTATGTGATGTGCTTCTTTGATGCAAGGTTCCAGAGGGACC	1320
85	1321	ATTCCAGTGTGACCTGGACACACAAGTCCGTGGACTTCT	1360
90	1361	TCAACATGATCGATAGCAAGAAGATCAGTCAACTTCCCTT	1400
95	1401	GGTGAAAGCCTACAAGCTGCAATCTGGTGCCTCCGTTGTC	1440

1441 GCAGGTCCCAGATTCACTGGAGGTGACATCATCCAGTGCA 1480  
 1481 CAGAGAACGGCAGCGCAGCTACTATCTACGTGACACCTGA 1520  
 1521 TGTGTCTTACTCTCAGAAGTACAGGGCAGCTATTTCATTAC 1560  
 1561 GCATCTACCAGCCAGATCACCTTCACACTCAGCTTGGATG 1600  
 1601 GAGCACCCCTTCAACCAGTATTACTTTGACAAGACCATCAA 1640  
 1641 CAAAGGTGACACTCTCACATACAATAGCTTCAACTTGGCA 1680  
 1681 AGTTTCAGCACACCATTGAACTCTCAGGCAACAATCTTC 1720  
 1721 AGATCGGCGTCAACCGGTCTCAGCGCGGAGACAAAGTCTA 1760  
 1761 CATCGACAAGATTGAGTTTCATCCAGTGAAC 1791

I. Un gène de structure qui code une protéine insecticide de *B.t. entomocidus* comportant la séquence :

1 ATGGAGGAGAGAACAACCAAAACCAATGCATTCCCTACAAC 40  
 5  
 41 GCTTGAGTAACCCAGAAGAGGTATTGCTTGATGGAGAACG 80  
 10  
 81 CATTTCAACCGGTAACCTCTTCCATCGACATCTCCTTGTC 120  
 121 TTGGTCCAGTTTCTGGTCAGCAACTTCGTGCCAGGTGGTG 160  
 15  
 161 GGITCCTTGTGCGGACTAATTGACTTCGTCTGGGGTATCGT 200  
 201 TGGTCCATCTCAATGGGATGCATTCTGGTGCAAATTGAG 240  
 20  
 241 CAGTTGATCAACGAGAGGATCGCTGAGTTCGCCAGGAACG 280  
 25  
 281 CTGCCATCGCTAACTTGAAGGATTGGGCAATAACTTCAA 320  
 321 CATCTATGTGGAGGCCCTTCAAAGAGTGGGAAGAGGACCC 360  
 30  
 361 AACCAACCCAGAGACCCGCACTAGGGTGATCGACAGATTCA 400  
 401 GAATCTTGGACGGCCTCTTGGAGAGAGATATCCCATCCTT 440  
 35  
 441 CAGAATCTCTGGCTTCGAAGTTCCTCTCTTGTCGGTGAC 480

5

481 GGTCAAGCAGCTAATOTTGACCTCGCTATCCTTCGAGACA 520

521 GTGTCATCTTTGGGGAAAGGTGGGGATTGACCACTATCAA 560

10

561 CGTCAATGAGAATTACAACAGACTTATCAGGCACATTGAC 600

601 GAGTACGCCGACCCTGTGCTAACACCTACAACCGTGGCT 640

15

641 TGAACAATCTCCCTAAGTCTACTTATCAAGATTGGATTAC 680

681 CTACAACAGGTTGAGGAGAGACTTGACCCTCAGTTTTG 720

20

721 GACATGCGAGCTTCTTCCCGAACTATGACAACAGGAGAT 760

761 ACCCTATCCAACCACTGGGTCAACTTACCAGAGAAGTCTA 800

25

801 TACTGACCCACTTATCAACTTCAACCCTCAGTTGCAAAGT 840

841 GTCGCCCAACTTCCCACATTCAACGTGATGGAGTCCAGCC 880

30

881 GTATCAGGAACCCACACTTGTTTGACATCTGAACAACCT 920

35

921 TACTATCTTCACCGATTGGTTGAGCGTTGGGCGTAACTTC 960

961 TATTGGGGTGGACACAGGGTCACTCTCCTCTCTATTGGAG 1000

40

1001 GTGGGAACATTACCTCTCCTATCTATGGACGTGAGGCAAA 1040

1041 CCAGGAGCCACCACGTAGTTTCACCTTCAACGGTCCAGTC 1080

45

1081 TTCAGAACCTTGTCTAACCTTACCTTGAGATTGCTCCAGC 1120

50

1121 AACCTTGGCCAGCTCCACCTTTCAACCTTAGAGGTGTTGA 1160

1161 GGGCGTTGAGTTCTCTACTCCTACCAACTCCTTCAGTTAC 1200

55

1201 AGAGGTAGAGGAACCGTTGATTCCCTGACCAGAACTCCCAC 1240

5	1241	CAGAGGACAATAGCGTGCCACCCAGGGGAGGCTACTCCCA	1280
	1281	CAGGTTGTGCCACGCAACCTTCGTGCAGCGTTCCGGAAC	1320
10	1321	CCATTCCCTCACTACAGGAGTTGTGTTCTCATGGACTGATC	1360
	1361	GTAGTGCTACTCTCACTAATACCATTGATCCCGAGAGGAT	1400
15	1401	CAATCAAATCCCATTGGTCAAGGGTTCCGTGTGTGGGA	1440
	1441	GGAAGTTCTGTTCATCACAGGACCGGCTTCACAGGAGGTG	1480
20	1481	ATATTCTTAGAAGAAACACTTTTGGCGACTTTGTGAGCCT	1520
	1521	CCAAGTTAACATCAACTCTCCAATTACTCAAAGATATCGT	1560
25	1561	CTCAGGTTTCGTTACGCATCTTCCCGTGACGCTAGAGTCA	1600
30	1601	TCGTGCTCACCGGAGCAGCTTCTACCGGTGTCGGTGGACA	1640
	1641	AGTCTCCGTGAACATGCCACTCCAGAAGACTATGGAGATC	1680
35	1681	GGCGAGAACTTGACATCCAGGACCTTCAGATACACCGACT	1720
	1721	TCTCTAACCCCTTTCAGTTTCCGTGCCAACCCCTGACATCAT	1760
40	1761	TGGCATTAGCGAACAACCTCTCTTTGGAGCTGGTAGCATC	1800
	1801	TCATCTGGCGAATTGTACATTGACAAGATTGAGATCATTC	1840
45	1841	TTGCCGACGCTACCTTCGAGGCTGAGTCTGACCTTGAGAG	1880
	1881	AGCCCCAGAAGGCTGTGAACGCCCTCTTTACCTCCTCTAAT	1920
50	1921	CAGATTGGCTTGAAAAGTACGCTTACTGACTATCACATTG	1960
55	1961	ACCAAGTGTCCAACCTGGTCTGACTGCCCTTAGCGATGAGTT	2000

5

2001 CTGCCCTCGACGAGAAGCGTGAACCTCTCCGAGAAGTTAAA 2040

10

2041 CACGCCAAGCGTCTCAGCGACGAGAGGAATCTCTTGCAAG 2080

2081 ACCCCCACTTCAGAGGCATCAACAGGCAGCCAGACCGTGG 2120

15

2121 TTGGAGAGGAAGCACCAGCATCACCATCCAAAGGAGGCGAC 2160

2161 GATGTGTTCAAGGAGAAGTACGTACCCCTCCCAGGAATG 2200

20

2201 TGGACGAGTGCTACCCCTACCTACTTGTACCAGAAGATCGA 2240

2241 TGAGTCCAAACTCAAAGCCTACACCAGGTATGAACCTAGA 2280

25

2281 GGCTACATCGAAGACAGCCAAGACCTTGAAATCTACCTCA 2320

2321 TCAGGTACAATGCCAAGCACGAGATCGTGAATGTCCCAGG 2360

30

2361 TACTGGTTCCCTCTGGCCACTTTCTGCCCAAATGCCCAAT 2400

2401 GGGAAAGTGTGGAGAGCCTAACAGATGCGCTCCACACCTTG 2440

35

2441 AGTGGAAATCCTGACTTGGACTGCTCCTGCAGGGATGGCGA 2480

40

2481 GAAGTGTGCCCACCAATCTCATCACTTCACCTTGGACATC 2520

2521 GATGTGGGATGTACTGACCTGAATGAGGACCTCGGAGTCT 2560

45

2561 GGGTCATCTTCAAGATCAAGACCCCAAGACGGACACGCAAG 2600

2601 ACTTGGCAACCTTGAGTTTCTCGAAGAGAAACCATTGCTC 2640

50

2641 GGTGAAGCTCTCGCTCGTGTGAAGAGAGCAGAGAAGAAGT 2680

2681 GGAGGGACAAACGTGAGAACTCCAACTCGAGACTAACAT 2720

55

2721 CGTTTACAAGGAGGCCAAGAGTCCGTGGATGCTTTGTTTC 2760



5	2761	GTGAACCTCCCAATATGATAGGTTGCAAGTGGACACCAACA	2800
10	2801	TCGCCATGATCCACGCTGCAGACAAACGTGTGCACAGGAT	2840
15	2841	TCGTGAGGCTTACTTGCCCTGAGTTGTCCGTGATCCCTGGT	2880
20	2881	GTGAACGCTGCCATCTTCGAGGAAC TTGAGGGACGTATCT	2920
25	2921	TTACCGCATACTCCTTGACGATGCCAGAAACGTCATCAA	2960
30	2961	GAACGGTGACTTCAACAATGGCCTCTTGTGCTGGAATGTG	3000
35	3001	AAAGGTCATGTGGACGTGGAGGAACAGAACAAATCACC GTT	3040
40	3041	CCGTCCTGGTTATCCCTGAGTGGGAAGCTGAAGTGTCCCA	3080
45	3081	AGAGGTTAGAGTCTGTCCAGGTAGAGGCTACATTCTCCGT	3120
50	3121	GTGACCGCTTACAAGGAGGGATACGGTGAGGGTTGCGTGA	3160
55	3161	CCATCCACGAGATCGAGGACAAACCCGACGAGCTTAAGTT	3200
60	3201	CTCCAACTGCGTCGAGGAAGAAGTCTATCCCAACAACACC	3240
65	3241	GTTACTTGCAACAATACTACACTGGGACCCAGGAAGAGTACG	3280
70	3281	AAGGTACCTACACTAGCCGTAACCAAGGTTACGACSAAGC	3320
75	3321	TTACGGAAACAATCCTTCGGTTCTCTGCTGACTATGCCTCC	3360
80	3361	GTGTACGAGGAGAAATCCTACACAGATGGCAGACGTGAGA	3400
85	3401	ACCCCTGCGAGTCCAACAGAGGTTACGGTGACTACACACC	3440
90	3441	ACTTCCAGCAGGCTATGTTACCAAGGACCTTGAGTACTTT	3480
95	3481	CCTGAGACCGACAAAGTGTGGATCGAGATCGGGTAAACCG	3520

3521 AGGGAACCTTCATCGGTGGACAGCGTGGAGCTTCTCTTGAT 3560

3561 GGAGGAA 3567.

J. Un gène de structure qui code - une protéine insecticide P2 comportant la séquence :

1 ATGGACAAACAACGTCCTTGAACCTCTGGTAGAACCAACATCT 40  
 5  
 41 GCGACGCATACAACGTCGTTGGCTCAGGATCCATTTCAGCTT 80  
 10  
 81 CGAACACAAGAGCCTCGACACTATTTCAGAGGAGTGGATG 120  
 121 GAATGGAAACGTAAGTACCACTCTCTCTACGTCGCACCTG 160  
 15  
 161 TGGTTGGAACAGTGTCCAGCTTCCTTCTCAAGAAGGTGGG 200  
 201 CTCTCTCATCGGAAACGTATCTTGTCCGAACCTCTGGGGT 240  
 20  
 241 ATCATCTTTCCATCTGGGTCCACTAATCTCATGCAAGACA 280  
 25  
 281 TCTTGAGGGAGACCGAACAGTTTCTCAACCAGCCTCTCAA 320  
 321 CACTGATACCTTGGGTAGAGTCAACGCTGAGTTGATCGGT 360  
 30  
 361 CTCCAAGCAAACATTCTGTGAGTTCAACCAGCAAGTGGACA 400  
 35  
 401 ACTTCTTGAATCCAACTCAGAATCCTGTGCCTCTTTCCAT 440  
 441 CACTTCTTCCGTGAACACTATGCAGCAACTCTTCTCTCAAC 480  
 40  
 481 AGATTGCCTCAGTTTCAGATTCAAGGCTACCAAGTTGCTCC 520  
 521 TTCTTCCACTCTTTGCTCAGGCTGCCAACATGCACTTGTG 560  
 45  
 561 CTTTCATACGTGACGTGATCCTCAACGCTGACGAATGGGGA 600  
 50  
 55

5

601 ATCTCTGCGGCCACTCTTAGGACATACAGAGACTACTGA 640

641 GGAACACACTCGTGTACTTCCAACTATTGCATCAACAC 680

10

681 TTATCAGACTGCCTTTCGTGGACTCAATACTAGGCTTCAC 720

721 GACATGCTTGAGTTCAGGACCTACATGTTCTTAACTGT 760

15

761 TTGAGTACGTCAGCATTGGAGTCTCTTCAAGTACCAGAG 800

20

801 CTTGATGGTGTCTCTCTGGAGCCAACTCTACGCCTCTGGC 840

841 AGTGGACCACAGCAAACCTCAGAGCTTCACAGCTCAGAACT 880

25

881 GGCCATTCTTTGATAGCTTGTTCCAAGTCAACTCCAATA 920

921 CATTCTCAGTGGTATCTCTGGGACCAGACTCTCCATAACC 960

30

961 TTTCCCAACATTGGTGGACTTCCAGGCTCCACTACAACCC 1000

1001 ATAGCCTTAACTCTGCCAGAGTGAACTACAGTGGAGGTGT 1040

35

1041 CAGCTCTGGATTGATTGGTGCAACTAAGTTGAACCACAAC 1080

1081 TTCAATTGCTCCACCGTCTTGCCACCTCTGAGCACACCGT 1120

40

1121 TTGTGAGGTCCTGGCTTGACAGCGGTACTGATCGCGAAGG 1160

45

1161 AGTTGCTACCTCTACAAACTGGCAAACCGAGTCTTCCAA 1200

1201 ACCACTCTTAGCCTTCGGTGTGGAGCTTCTCTGCACGTG 1240

50

1241 GGAATTCAAACACTACTTCCAGACTACTTCATTAGGAACAT 1280

1281 CTCTGGTGTCTCTCTCGTCAATCAGGAATGAAGACCTCACC 1320

55

1321 CGTCCACTTCATTACACCAGATTAGGAACATCGAGTCTC 1360

1361 CATCCGCTACTCCAGGAGGTGCRAGAGCTTACCTCGTGTC 1400  
 1401 TGTCCATAACAGGAAGAACAACATCTACGCTGCCAACGAG 1440  
 1441 AATGGCACCATGATTACCTTGCACCCAGAAGATTACACTG 1480  
 1481 GATTCAACCATCTCTCCAAATCCATGCTACCCAAAGTGAACAA 1520  
 1521 TCAGACACGCACCTTCATCTCCGAAAAGTTCGGAAATCAA 1560  
 1561 GGTGACTCCTTGAGGTTTCGAGCAATCCAACACTACCGCTA 1600  
 1601 GGTACACTTTGAGAGGCAATGGAAACAGCTACAACCTTTA 1640  
 1641 CTTGAGAGTTAGCTCCATTGGTAACTCCACCATCCGTGTT 1680  
 1681 ACCATCAACGGACGTGTTTACACAGTCTCTAATGTGAACA 1720  
 1721 CTACACGAACAATGATGGCGTTAACGACAACGGAGCCAG 1760  
 1761 ATTCAGCGACATCAACATTGGCAACATCGTGGCCTCTGAC 1800  
 1801 AACACTAACGTTACTTTGGACATCAATGTGACCCCTCAATT 1840  
 1841 CTGGAACCTCCATTGATCTCATGAACATCATGTTTGTGCC 1880  
 1881 AACTAACCTCCCTCCATTGTAC 1902 ou

K. Une séquence de gène de structure codant une protéine de fusion comprenant les acides aminés 610 N-terminaux de *B.t.k.* HD-1 et les acides aminés 567 C-terminaux de *B.t.k.* HD-73, ledit gène comportant la séquence :

1 ATGGGACAAACAACCCAAACATCAACGAATGCATTCCATACA 40  
 41 ACTGOTTGAGTAAACCCAGAAGTTGAAGTACTTGGTGGAGA 80  
 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120  
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160  
 161 CTGGGTTTCGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200  
 201 CTTTGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATT 240  
 241 GAGCAGTTGATCAACCCAGAGGATCGAAGAGTTCGCCAGGA 280  
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320  
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360  
 361 CCTACTAACCAGCTCTCCGCCAGGAAATGCGTATTCAAT 400  
 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCAT 440

5	441	GTTCCGAGTCCAGAACTACCAAGTTCCTCTCTTTGTCCGTG	480
	481	TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG	520
10	521	ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTTCGATGCTGC	560
	561	AACCTCAATAGCCGTTACAACGACCTTACTAGGCTGATT	600
15	601	GGAAACTACACCGACCACGCTGTTGTTGGTACAACACTG	640
	641	GCTTGGAGCGTGTCTGGGTCCTGATTCTAGAGATTGGAT	680
20	681	TAGATACAACCAAGTTCAGGAGAGAATTGACCCCTCACAGTT	720
	721	TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA	760
25	761	CCTACCCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT	800
30			
35			
40			
45			
50			
55			

5  
 801 CTATACTAAGCCAGTTCTTGAGAACTTCGACGGTAGCTTC 840  
 841 CGTGGTTCGCCCCAAGGTATCGAAGGCTCCATCAGGAGCC 880  
 10  
 881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920  
 921 CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG 960  
 15  
 961 ATCATGGCCTCTCCAGTTGGATTGAGCGGGCCCGAGTTTA 1000  
 1001 CCTTTCCTCTCTATGGAACATATGGGAAACGCCGCTCCACA 1040  
 1041 ACAACGTATCGTTGCTCAACTAGGTCAGGCTGTCTACAGA 1080  
 25  
 1081 ACCTTGTCTTCCACCTTGTACAGAAGACCCCTTCAATATCG 1120  
 1121 GTATCAACAACCAGCAACTTTCGGTTCCTTGACGGAACAGA 1160  
 30  
 1161 GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT 1200  
 1201 TACAGAAAGAGCGGAACCGTTGATTCCCTTGGACCAAATCC 1240  
 1241 CACCACAGAACAAATGTGCCACCCAGGCAAGGATTCTC 1280  
 40  
 1281 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTCC 1320  
 1321 AGCAACAGTTCGGTGAGCATCATCAGAGCTCCTATGTTCT 1360  
 45  
 1361 CATGGATTTCATCGTAGTGCTGAGTTCAACAATATCATTC 1400  
 1401 TTCCTCTCAAATCACCCTAAATCCCATTTGACCAAGTCTACT 1440  
 1441 AACCTTGGATCTGGAACCTCTGTGCTGAAAGGACCAGGCT 1480  
 55  
 1481 TCACAGGAGGTGATATTCTTAGAAGAACTCTCTCGGCCCA 1520



5

1521 GATTAGCAGCCTCAGAGTTAACAATCACTGCACCACCTTTCT 1560

1561 CAAAGATATCGTGTGAGGATTGTTACGCATCTACCACTA 1600

10

1601 ACTTGCAATTCCACACCTCCATCGACGGAAGGCCTATCAA 1640

1641 TCAGGGTAACTTCTCCGCAACCATGTCAAGCGGCAGCAAC 1680

15

1681 TTGCAATCCGGCAGCTTCAGAACCGTCGGTTTCACTACTC 1720

1721 CTTTCAACTTCTCTAACGGATCAAGCGTTTTCACCCCTTAG 1760

20

1761 CGCTCATGTGTTCAATTCTGGCAATGAAGGTACATTGAC 1800

25

1801 CGTATTGAGTTTGTGCCTGCCGAAGTTACCTCGAGGCTG 1840

1841 AGTACAACCTTGAGAGAGCCGAGAAGGCTGTGAACGCCCT 1880

30

1881 CTTTACCTCCACCAATCAGCTTGGCTTGAAACTTACGTT 1920

1921 ACTGACTATCACATTGACCAAGTGTCCTCAACTTGGTCACCT 1960

35

1961 ACCTTAGCGATGAGTTCTGCCTCGACGAGAAGCTGAACT 2000

2001 CTCCGAGAAAGTTAAACACGCCAAGCGTCTCAGCGACGAG 2040

40

2041 AGGAATCTCTTGCAAGACTCCAAGTTCAAAGACATCAACA 2080

2081 GGCAGCCAGAACGTGTTGGGTTGGAAGCACCGGGATCAC 2120

45

2121 CATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTACGTC 2160

50

2161 ACCCTCTCCGGAACCTTCGACGAGTGCTACCTACCTACT 2200

2201 TGTACCAGAAGATCGATGAGTCCAAACTCAAAGCCTTCAC 2240

55

2241 CAGGTATCAACTTAGAGGCTACATCGAAGACAGCCCAAGAC 2280

5

2281 CTTGAAATCTACTCGATCAGGTACAATGCCAAGCACCAGA 2320

2321 CCGTGAATGTCCCAGGTACTGGTTCCTCTGCCCACTTTC 2360

10

2361 TGCCCAATCTCCCATTTGGGAAGTGTGGAGAGCCCTAACAGA 2400

2401 TGCCTCCACACCTTGAGTGGAACTCCTGACTTGGACTGCT 2440

15

2441 CCTGCAGGGATGGCGAGAAGTGTGCCCACTTCTCATCA 2480

2481 CTTCTCCTTGGACATCGATGTGGGATGTACTGACCTGAAT 2520

20

2521 GAGGACCTCGGAGTCTGGGTCTCTTCAAGATCAAGACCC 2560

2561 AAGACCGACACGCAAGACTTGGCAACCTTGAGTTTCTCGA 2600

25

2601 AGAGAAACCATTTGGTCGGTGAAGCTCTCGCTCGTGTGAAG 2640

30

2641 AGAGCAGAGAAGAAGTGGAGSGACAAACGTGAGAACTCG 2680

2681 AATGGGAAGCTAACATCGTTTACAAGGAGGCCAAGAGTC 2720

35

2721 CGTGGATGCTTTGTTTCGTGAACCTCCCAATATGATCAGTTG 2760

2761 CAAGCCGACACCAACATCGCCATGATCCACGCCGCAGACA 2800

40

2801 AACGTGTGCACAGCATTTCGTGAGGCTTACTTGCTGAGTT 2840

45

2841 GTCCGTGATCCCTGGTGTGAACGCTGCCATCTTCGAGGAA 2880

2881 CTTGAGGGACGTATCTTTACCGCATCTCCTTTGACGATG 2920

50

2921 CCAGAAACGTGATCAAGAACGGTGACTTCAACAAATGGCCT 2960

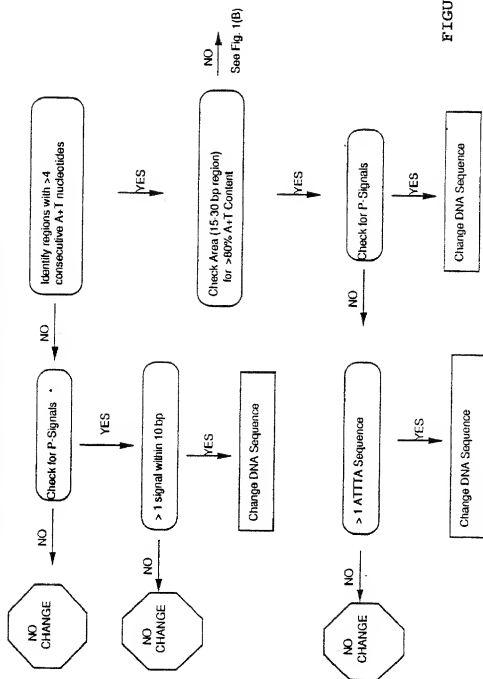
2961 CAGCTGCTGGAAATGTGAAGGTCATGTGGACGTGGAGGAA 3000

55

3001 CAGAACAAATCAGCGTTCCGTCTCGTTGTGCTGAGTGGG 3040

3041 AAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGGTAG 3080  
 3081 AGGCTACATTCTCCGTGTGACCGCTTACAGGAGGGATAC 3120  
 3121 GGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAACA 3160  
 3161 CCGACGAGCTTAAGTTCTCCAAGTCCGTCGAGGAAGAAAT 3200  
 3201 CTATCCCAACAACACCGTTACTTGCAACGACTACACTGTG 3240  
 3241 AATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTAACA 3280  
 3281 GAGGTTACAACGAAGCTCCTTCCGTTCCCTGCTGACTATGC 3320  
 3321 CTCCTGTACGAGGAGAAATCCTACACAGATGGCAGACGT 3360  
 3361 GAGAACCCTTGCGAGTTCAACAGAGGTTACAGGGACTACA 3400  
 3401 CACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGAGTA 3440  
 3441 CTTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAA 3480  
 3481 ACCGAGGGAACCTTCAATCGTGGACAGCGTGGAGCTTCTCT 3520  
 3521 TGATGGAGGAA 3531.

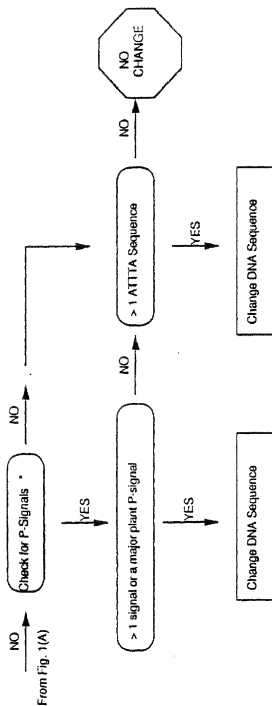
Determination of DNA regions in genes to modify by  
site-directed mutagenesis for increased expression in plants



\* Polyadenylation Signal Sequences

FIGURE 1 (A)

Determination of DNA regions in genes to modify by  
allele-directed mutagenesis for increased expression in plants



\* Polyadenylation Signal Sequences

FIGURE 1 (B)

1 ATGGCTATAGAACTGGTTACACCCCAATCGATATTTCTCT 40  
 41 TGTGCGCTAACGCAATTTCTTTTGTAGTGAATTTGTTCCCGG 80  
 81 TGCTGGATTGTGTAGGACTAGTTGATATAATATGGGGA 120  
   T C  
 121 ATTTTGGTCCCTCTCAATGGGACGCATTCTTGTACAAA 160  
 161 TTGAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAG 200  
                   C C C G C G  
 201 GAACCAAGCCATTCTTAGATTAGAAGGACTAAGCAATCTT 240  
                   T  
 241 TATCAAATTTACGCAGAAATCTTTTAGAGAGTGGGAAGCAG 280  
 281 ATCCTACTAATCCAGCATTAAAGAGAAGAGATCGGTATTCA 320  
 321 ATTCAATGACATGAACAGTGCCCTTACAACCGCTATTCCT 360  
 361 CTTTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAG 400  
   CC C C  
 401 TATATGTTCAAGCTGCAAATTTACATTTATCAGTTTGTAG 440  
           G C C CC C CC C  
 441 AGATGTTTCAGTGTTTGGACAAAGGTGGGGAATTTGATGCC 480  
 481 GCGACTATCAATAGTCGTTATTAATGATTTAACTAGGCTTA 520  
 521 TTGGCAACTATACAGATCATGCTGTACGCTGGTACAATAC 560  
 561 GGGATTAGAGCGTGTATGGGGACCGGATTCTAGAGATTGG 600  
 601 ATAAGATATAATCAATTTAGAAGAGAATTAACACTAACTG 640  
           C G C C G C GC T  
 641 TATTAGATATCGITTCTCTATTTCCGAACTATGATAGTAG 680  
 681 AACGTATCCATTCGAACAGTTTCCCAATTAACAAGAGAA 720

FIGURE 2A

721	ATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTT	760
761	TCGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAG	800
801	TCCACATTTGATGGATATACCTTAATAGTATAACCATCTAT	840
841	ACGGATGCTCATAGAGGAGAAATATTATTGGTCAGGGCATC C C C T C	880
881	AAATAATGGCTTCTCCTGTAGGGTTTTCGGGGCCAGAATT G C	920
921	CACTTTTCCGCTATATGGAACCTATGGGAAATGCAGCTCCA	960
961	CAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATA	1000
1001	GAACATTATCGTCCACCTTATATAGAAGACCTTTTAAATAT C	1040
1041	AGGGATAAATAATCAACAACCTATCTGTTCTTGACGGGACA C C C C	1080
1081	GAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTG	1120
1121	TATACAGAAAAGCGGAACGGTAGATTCTGCTGGATGAAAT	1160
1161	ACCGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTT	1200
1201	AGTCATCGATTAAAGCCATGTTTCAATGTTTCGTTACGGCT	1240
1241	TTAGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTT	1280
1281	CTCTTGATACATCGTAGTGCTGAATTTAATAATATAATT G C C C C C	1320
1321	CCTTCATCACAAATTACACAAATACCTTTAACAAAATCTA C C C AC C C G	1360
1361	CTAATCTTGGCTCTGGAACCTTCTGTCGTTAAAGGACCAGG	1400

FIGURE 2B

1401	ATTACAGGAGGAGATATTCTTCGAAGAACTTCACCTGGC	1440
1441	CAGATTTCAACCTTAAGAGTAAATATTACTGCACCATTAT	1480
1481	CACAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCAC	1520
1521	AAATTTACAATTCCATACATCAATTGACGGAAGACCTATT	1560
	CC T G C	
1561	AATCAGGGGAATTTTTCAGCAACTATGAGTAGTGGGAGTA	1600
1601	ATTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTAC	1640
1641	TCCGTTTAACTTTTCAAATGGATCAAGTGTATTTACGTTA	1680
1681	AGTGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAG	1720
1721	ATCGAATTGAATTTGTTCCGGCA	1743

FIGURE 2C



1	ATGGATAACAATCCGACACATCAATGAATGCATTCCCTTATA	40
	C C A C A C	
41	ATTGTTTAAAGTAACCCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAAACTGGTTACACCCCAATCGATATTTCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTTCTTTTGAAGTGAATTTGTTCCCGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTTGGTCCCTCTCAATGGGACGCATTTCCTTGACAAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCCCTCT	440
	C C T G C A C AT	
441	TTTTCAGTTCAAATTTATCAAGTTCCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTGTAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTTTGGACAAAGGTGGGGATTGTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATcATGCTGTaCGCTGGTACAATACGG	640
	A C C CC C T T C T	
641	GATTAGAGCGTGTATGGGGACCGGATTCTAGAGATTGGAT	680
	C G C T T	

FIGURE 3A

681	AAGATATAATCAATTTAGAGAAGAGAAATTAACACTAAGTGA	720
	T C C G C G G C C A T	
721	TTAGATATCGTTTCTCTATTTCCGAACATGATAGTAGAA	760
	G C T G C C CTCC	
761	CGTATCCAATTCGAACAGTTTCCCAATTAACAAGAGAAAT	800
	C C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAATTTGATGGTAGTTTT	840
	C T TC T G C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAGTC	880
	T T T C A T C CTCC C C	
881	CACATTTGATGGATATACTTAATAGTATAACCATCTATAC	920
	C C CT G C C T C	
921	GGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATCAA	960
	C C G C T A C G	
961	ATAATGGCTTCTCCTGTAGGGTTTTTCGGGGCCAGAATTCA	1000
	C C A T A CAGC C G T	
1001	CTTTCCGCTATATGGAACATGCGGAATGCAGCTCCACA	1040
	C T C C C	
1041	ACAACGTATTTGTCTCAACTAGGTCAGGGCGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACCTTATATAGAAGACCTTTTAATATAG	1120
	C G T G C C C C	
1121	GGATAAATAATCAACAACCTATCTGTCTTGTACGGGACAGA	1160
	T C C C G T C A	
1161	ATTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTCTGCTGGATGAATAC	1240
	G C T CT C C	
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C CTC	
1281	TCATCGATTAAAGCCATGTTTCAATGTTTCGTTCAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	
1361	CTTGGATACATCGTAGTGTGAATTTAATAATATAATTCC	1400
	A T G C C C	

FIGURE 3B

1401	TTCATCACAAATTACACAAATACCTTTAAACAAAATCTACT	1440
	C T C C C A G C G	
1441	AATCTTGGCTCTGGAACCTCTGTCGTTAAAGACCAGGAT	1480
	C A G C	
1481	TTACAGGAGGAGATATTCTTTCGAAGAACTTCACCTGGCCA	1520
	C T A T	
1521	GATTTCAACCTTAAGAGTAAATATTACTGCACCATTATCA	1560
	AGC C C T C C C T T	
1561	CAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCACAA	1600
	T C G T A A	
1601	ATTTACAATTCCATACATCAATTGACGGAAGACCTATTAA	1640
	C G C C C C G C	
1641	TCAGGGGAATTTTTCAGCAACTATGAGTAGTGGGAGTAAT	1680
	T C C C C TCA C C C C	
1681	TTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTACTC	1720
	G A C C A C C C	
1721	CGTTTAACTTTTCAAATGGATCAAGTGTATTTACGTTAAG	1760
	T C C T C C T C C C T	
1761	TGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAGAT	1800
	C G T T G C T C	
1801	CGAATTGAATTTGTTCCGGCAGAAGTAACCTTTGAGGCAG	1840
	T G G T C T C T	
1841	AATAT	1845
	G C	

FIGURE 3C

1	ATGGATAACAATCCGAACATCAATGAATGCATTCCCTTATA	40
	C C A C A C	
41	ATTGTTTAAGTAACCCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAAACTGGTTACACCCCAATCGATATTTCCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTTCTTTTGGAGTGAATTTGTTCCCGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTGGTCCCTCTCAATGGGACGACATTTCCTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCRAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCTTACAACCGCTATTCCTCT	440
	C C T G C A C AT	
441	TTTTCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTGGAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTGTTGGACAAAGTGGGGATTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C C T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	
641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT	680
	C G G C T T A	

FIGURE 4A

681	AAGGTATAATCAATTTAGAAGAGAATTAACACTAACTGTA	720
	T A C C G C G G C C A T	
721	TTAGATATCGTTGCTCTGTTCCCGAATTATGATAGTAGAA	760
	G C T GT C C CTCC	
761	GATATCCAATTGGAACAGTTTCCCAATTAACAAGAGAAAT	800
	CC C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAAATTTGATGGTAGTTTT	840
	C T TC T G C C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC	880
	T T T C A T C G CTCC C C	
881	CACATTTGATGGATATACTTAACAGTATAACCATCTATAC	920
	C C CT G C T C	
921	GGATGCTCATAGGGGTTATTATTATTTGGTCAGGGCATCAA	960
	C C A AG G C T A C G	
961	ATAATGGCTTCTCCTCTAGGGTTTTTCGGGGCCAGAATCA	1000
	C C A T A CAGC C G T	
1001	CTTTTCCGCTATATGGAACATATGGGAAATGCAGCTCCACA	1040
	C T C C C	
1041	ACAACGTATTGTGCTCAACTAGGTCAGGGCGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTTAAATATAG	1120
	C G T C G C C C C	
1121	GGATAAATAATCAACAACATCTCTGTTCTTGACGGGACAGA	1160
	T C C C G T C A	
1161	ATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTGCTGGATGAAATAC	1240
	G C T CT C C	
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C CTC	
1281	TCATCGATTAAAGCCATGTTTCAATGTTTCGTTTCAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	
1361	CTTGGATACATCCGTAGTGCTGAATTTAATAATATAATTGC	1400
	C G C C C C C	

FIGURE 4B

1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTTAAATGGTTCTGTAAATTCAGGACCAGGATTTA	1480
	C C C C C	
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAAATAA	1520
	A C C C C C C	
1521	CATTGAGAATAGAGGGTATATTGAAGTTCCAATTCACCTC	1560
1561	CCATCGACATCTACCAGATATCGAGTTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAAACCCCGATTCACTCAACGTTAATTGGGGTAA	1640
	G T	
1641	TTCATCCATTTTTTCCAATACAGTACCAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTTGGTTATTTTG	1720
	C G C C C C	
1721	AAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTTGAATTTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGAAAAGAGCGCAGAAGGCGGTGAATGC	1880
	A TGCG	
1881	GCTGTTTACGTCTACAAACCAACTAGGGCTAAAAACAAT	1920
	CTGT ACGTCTACA C AGCT G ACTC G CA TG	
1921	G 1921	

FIGURE 4C

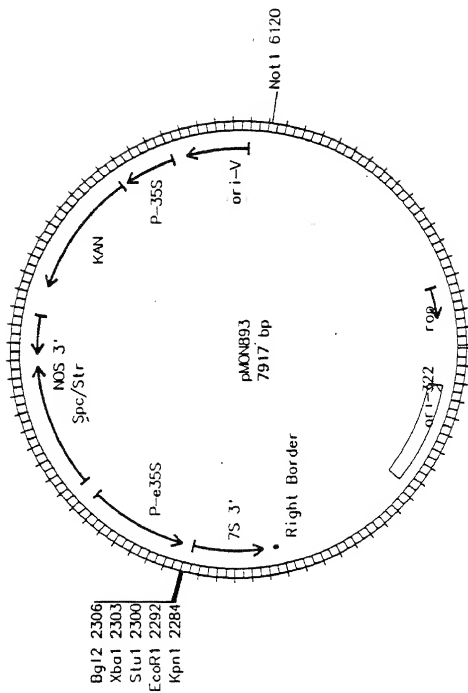


FIGURE 5

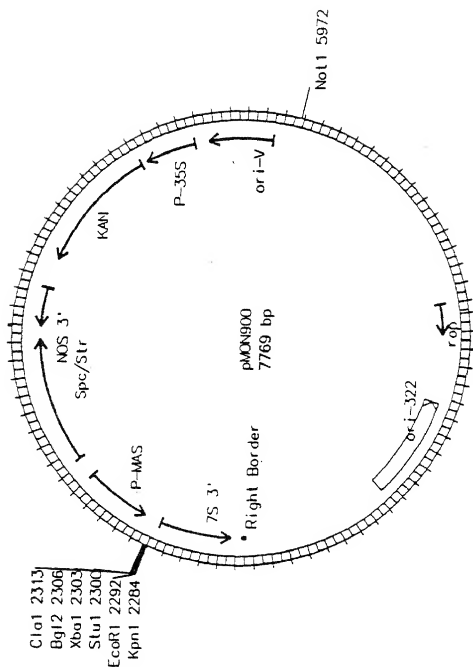
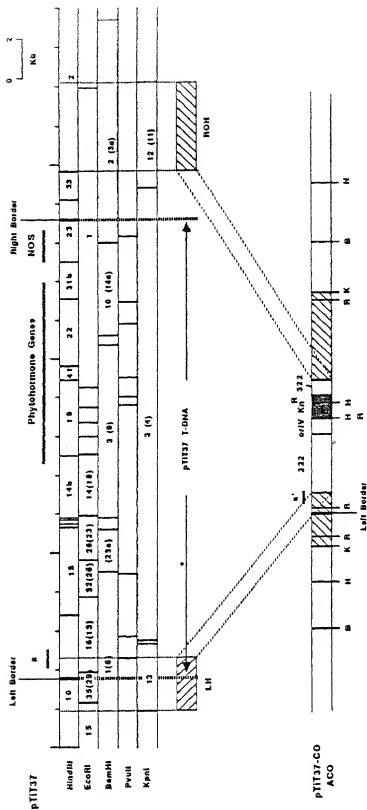


FIGURE 6





**FIGURE 7**

```

1  GAAAGAATAGAAACTGGTTACACCCCAATCGATATTTCCT  40
   ATGGCC T C T C C C
41  TGTGCTAACGCAATTTCTTTTGTAGTGAATTTGTTCCTCGG  80
   CT G A G GC C C G C G A
81  TGCTGGATTGTGTGTAGGACTAGTTGATATAATATGGGGA  120
   G C TC C C C C T
121  ATTTTGGTCCCTCTCAATGGGACGCATTCTTGTACAAA  160
   C A T C G G
161  TTGAACAGTTAATTAACCAAGAATAGAAGAATTCGCTAG  200
   G G C G G C G C
201  GAACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTT  240
   G C G G T G C
241  TATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAG  280
   C C T GAGC C C
281  ATCCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCA  320
   C TC CC C G A
321  ATTCAATGACATGAACAGTGGCCTTACAACCGCTATTTCCT  360
   C C T G C A C A
361  CTTTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAG  400
   T G C C G C C C G C
401  TATATGTTCAAGCTGCAAAATTACATTATCAGTTTGTAG  440
   G C A T C T CC CAGC GC TC
441  AGATGTTTCAGTGTGTTGGACAAAGGTGGGGATTGTGATGCC  480
   C AGC G C T
481  GCGACTATCAATAGTCGTTATAATGATTAACTAGGCTTA  520
   A C C C C CC T G
521  TTGGCAACTATACAGATTATGCTGTACGCTGGTACAATAC  560
   A C C CC C T T C
561  GGGATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGG  600
   T C G G C T T
601  GTAAGGTATAATCAATTTAGAAGAGAATTAACACTAACTG  640
   A T A C C G C G G C C A
641  TATTAGATATCGTTGCTCTGTTCGCCGAATTATGATAGTAG  680
   T G C T GT C C CTCC

```

FIGURE 8A

681	AAGATATCCCAATTCGAACAGTTTCCCAATTACAAGAGAA	720
	CC C T C T G C T C	
721	ATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTT	760
	C T TC T G C C C C	
761	TTCGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAG	800
	C T T T C A T C G CTCC C	
801	TCCACATTTGATGGATATACTTAACAGTATAACCATCTAT	840
	C C C CT G C T C	
841	ACGGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATC	880
	C C A AG G C T A C	
881	AAATAATGGCTTCTCCTGTAGGGTTTCGGGGCCAGAATT	920
	G C C A T A CAGC C G	
921	CACTTTCCGCTATATGGAACATATGGGAAATGCAGCTCCA	960
	T C T C C C	
961	CAACAACGTATTGTTGCTCAACTAGGTCAGGGCGGTGTATA	1000
	C T C C	
1001	GAACATTATCGTCCACTTTATATAGAAGACCTTTTAATAT	1040
	C G T C G C C C	
1041	AGGGATAAATAATCAACAACATCTGTCTTGACGGGACA	1080
	C T C C C G T C A	
1081	GAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTG	1120
	G C C T T C	
1121	TATACAGAAAAAGCGGAACGGTAGATTTCGCTGGATGAAAT	1160
	T G C T CT C	
1161	ACCGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTT	1200
	C A C T C C	
1201	AGTCATCGATTAAAGCCATGTTTCAATGTTTCGTTTCAGGCT	1240
	TCC CA G G C G C C C A	
1241	TTAGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTT	1280
	C C C TCC G C C C	
1281	CTCTTGATACATCGTAGTGCTGAATTTAATAATATAATT	1320
	C G C C C C C	
1321	GCATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAA	1360
	C	
1361	ACTTTCCTTTTAATGGTTCTGTAAATTCAGGACCAGGATT	1400
	C C C C	

FIGURE 8B

1401	TACTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAAAT	1440
	C            A    C   C            C   C   C   C	
1441	AACATTGAGAAATAGAGGGTATATTGAAGTTCCAATTCAC	1480
1481	TCCCATCGACATCTACCAGATATCGAGTTCGTGTACGGTA	1520
	C    A    GA	
1521	TGCTTCTGTAAACCCCGATTACCTCAACGTTAATTGGGGT	1560
	G    T	
1561	AATTCATCCATTTTTTCCAATACAGTACCAGCTACAGCTA	1600
	C   C    T	
1601	CGTCATTAGATAATCTACAATCAAGTGATTTGGTTATTT	1640
	C   C   G    C   C   C    C   C	
1641	TGAAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATA	1680
	C   C    C   C    C   C	
1681	GTAGGTGTTAGAAATTTTAGTGGGACTGCAGGAGTGATAA	1720
	G    C    T	
1721	TAGACAGATTTGAATTTATTCAGTTACTGCAACACTCGA	1760
	C    C   G   C	
1761	GGCTGAA	1767
	G	

FIGURE 8C

1	ATGGATAACAATCCGAACATCAATGAATGCATTCCTTATA	40
	C C A C A C	
41	ATTGTTTAAGTAACCCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAACTGGTTACACCCCAATCGATATTTCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTCTCTTTTGTGAGTGAATTTGTTCCCGGTG	160
	CT G A G G C C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTTGGTCCCTCTCAATGGGACGCATTTCCTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCCTCT	440
	C C T G C A C AT	
441	TTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAAATTTACATTTATCAGTTTGGAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTGTTGGACAAGGTGGGGATTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	
641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT	680
	C G G C T T A	

FIGURE 9A

581	ARGGTACAAATTC	ATTAGAGAGAATT	TAACACTAACTGTA	720
	T A C C G C G		G C C A T	
721	TTAGATATCGTTG	CTCTGTCCCGAATT	ATGATAGTAGAA	760
	G C T GT	C	CTCC	
761	GATATCCAATTCGA	ACAGTTTCCCAATTA	ACAAGAGAAAT	800
	CC C T C T	G	C T C	
801	TTATACAAACCCG	ATTATAGAAAAATTT	GATGGTAGT	840
	C T	TC T G C C C	C C	
841	CGAGGCTCGGCTC	AGGCGATAGAAA	AGATATTAGGAGTC	880
	T T T C A T C	G CTCC	C C	
881	CACATTTGATGGAT	ATACTTAAACAGTATA	AACCATCTATAC	920
	C	C CT G	C T C	
921	GGATGCTCATAGGG	GTTATTATTATGGT	CAGGGCATCAA	960
	C A AG G	C	T A C G	
961	ATAATGGCTTCTC	CTGTAGGGTTTTC	GGGCGCAGAATCA	1000
	C C	A T A CAGC	C G T	
1001	CTTTCCGCTATAT	GGAATATGGGAAATG	CAGCTCCACA	1040
	C T C	C C		
1041	ACAACGTATTGTT	GCTCAACTAGGTC	CAGGGCGTGATAGA	1080
	C		T C C	
1081	ACATTATCGTCCAC	TTTATATAGAAGAC	CTTTTAATATAG	1120
	C G T	C G C	C C C	
1121	GGATAAATAATCA	AACTATCTGTTCTT	GACGGGACAGA	1160
	T C C C G	T C	A	
1161	ATTTGCTTATGGA	ACCTCCTCAAATTT	GCCATCCGCTGTA	1200
	G C C	T T C	T	
1201	TACAGAAAAAGCG	GAACGGTAGATT	TCGCTGGATGAAATAC	1240
	G	C T	CT C C	
1241	CGCCACAGAATAA	CAACGTGCCACCT	AGGCAAGGATTAG	1280
	A C	T C	CTC	
1281	TCATCGATTAAAG	CCATGTTTCAATG	TTTCGTTCAGGCTTT	1320
	C CA G G	C G C	C C A C	
1321	AGTAATAGTAGTG	TAAGTATAATAAG	AGSCTCCTATGTTCT	1360
	C C TCC	G C C C		
1361	CTTGGATACATCG	TAGTGTGAATTTA	ATAATAATTGC	1400
	C	G C C C C		

FIGURE 9B

1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTTAAATGGTTCTGTAATTTTCAGGACCAGGATTTA	1480
	C C C C C	
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAAATAA	1520
	A C C C C C C	
1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACCTC	1560
1561	CCATCGACATCTACCAGATATCGAGTTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAACCCCGATTCACTCAACGTTAATTGGGGTAA	1640
	G T	
1641	TTCATCCATTTTTTCCAATACAGTACCAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTTGGTTATTTTG	1720
	C G C C C C C	
1721	AAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATAGT	1760
	C C C C C	
1761	AGGTGTTAGAAAATTTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTGGAATTTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
1881	GCTGTTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT	1920
1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTTAGTTA	1960
1961	CGTATTTATCGGATGAATTTTGTCTGGATGAAAAGCGAGA	2000
2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGGACTCAGTGAT	2040
2041	GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTA	2080
2081	ATAGGCAACCAGAACSTGGGTGGGGCGGAAGTACAGGGAT	2120

FIGURE 9C

2121	TACCATCCAAGCAGGGGATGACGTATTTAAAGAAATTAC	2160
2161	GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT	2200
2201	ATTTGTATCAAAAAATCGATGAATCAAAATTAAAAGCCTT	2240
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA	2280
2281	GACTTAGAAATCTATTTAATTCGCTACAATGCAAAACATG	2320
2321	AAACAGTAAATGTGCCAGGTACGGGTTCCTTATGGCCGCT	2360
2361	TTCAGCCCAAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
2441	GTTCTGTAGGGATGGAGAAAAGTGTGCCCATCATTCGCA	2480
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
2521	AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAGA	2560
2561	CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT	2600
2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
2641	AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAAAAAT	2680
2681	TGGAATGGGAAACAAATATCGTTTATAAAGAGGCAAAAGA	2720
2721	ATCTGTAGATGCTTTATTTGTAAACTCTCAATATGATCAA	2760
2761	TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG	2800
2801	ATAAACGTGTTTCATAGCATTCGAGAAGCTTATCTGCCTGA	2840

FIGURE 9D



2841	GCTGTCGTGATTCGGGTGTCAATGCGGCTATTTTGA	2880
2881	GAATTAGAAGGGCGTATTTTCACTGCATTCCTCCCTATATG	2920
2921	ATGCGAGAAATGTCATTAAAAATGGTGATTTTAATAATGG	2960
2961	CTTATCCTGCTGGAACGTGAAAGGCATGTAGATGTAGAA	3000
3001	GAACAAAACAACCAACGTTCCGTCCTTGTTGTTCCGGAAT	3040
3041	GGGAAGCAGAAGTGTCAACAAGAAGTTCGTGTCTGTCCGGG	3080
3081	TCGTGGCTATATCCTTCGTGTCACAGCGTACAAGGAGGGA	3120
3121	TATGGAGAAGGTTGCGTAACCATTCATGAGATCGAGAACA	3160
3161	ATACAGACGAACTGAAGTTTAGCAACTGCGTAGAAGAGGA	3200
3201	AATCTATCCAATAACACGGTAACGTGTAATGATTATACT	3240
3241	GTAATCAAGAAGAAATACGGAGGTGCGTACACTTCTCGTA	3280
3281	ATCGAGGATATAACGAAGTCCTTCCGTACCAGCTGATTA	3320
3321	TGCGTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA	3360
3361	AGAGAGAATCCTTGTGAATTTAACAGAGGTATAGGGATT	3400
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
3481	GAAACGGAAGGAACATTTATCGTGGACAGCGTGGAATTAC	3520
3521	TCCTTATGGAGGAA	3534

FIGURE 9E

1	ATGGATAACAATCCGAACATCAATGAATGCATTCCCTTATA	40
	C C A A C	
41	ATTGTTTAAGTAACCCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAACTGGTTACACCCCAATCGATATTTCCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTTCTTTTGAGTGAATTGTTCCTCCGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTTGGTCCCTCTCAATGGGACGCATTTCTTGTAACAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAGAATAGAAGAATTCGCTAGGA	280
	G G C G C C	
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCCCTCT	440
	C C T G C A C AT	
441	TTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTGTAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTTTGGACAAGGTGGGGATTGTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	

FIGURE 10A

641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT	680
	C G G C T T A	
681	AAGGTATAATCAATTTAGAAAGAGAATTAACTACTGTAT	720
	T A C C G C G G C C A T	
721	TTAGATATCGTTGCTCTGTTCCCGAATTATGATAGTAGAA	760
	G C T GT C C CTCC	
761	GATATCCAATTCGAACAGTTTCCCAATTAACAAGAGAAAT	800
	CC C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTT	840
	C T T C T G C C C C C	
841	CSAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC	880
	T T T C A T C G CTCC C C	
881	CACATTTGATGGATATACTTAACAGTATAACCATCTATAC	920
	C C CT G C T C	
921	GGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATCAA	960
	C C A AG G C T A C G	
961	ATAATGGCTTCTCCTGTAGGGTTTTTCGGGCCAGAATTCA	1000
	C C A T A CAGC C G T	
1001	CTTTTCCGCTATATGGAACATATGGGAAATGCAGCTCCACA	1040
	C T C C C	
1041	ACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAG	1120
	C G T C G C C C C	
1121	GGATAAATAATCAACAACCTATCTGTTCTTGACGGGACAGA	1160
	T C C C G T C A	
1161	ATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAATAC	1240
	G C T CT C C	
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C CTC	
1281	TCATCGATTAGCCATGTTTCAATGTTTCGTTTCAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	

FIGURE 10B

1361	CTTGGATACATCGTAGTGCCTGAATTTAATAATATAATTGC	1400
	C G C C C C C	
1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTTAAATGGTCTCTGTAATTTTCAGGACCAGGATTTA	1480
	C C C C	
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAAATAA	1520
	A C C C C C C	
1521	CATTTCAGAAATAGAGGGTATATTGAAGTTCCAATTCACCTTC	1560
1561	CCATCGACATCTACCAGATATCGAGTTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAAACCCCGATTCACTCAACGTTAAATTGGGGTAA	1640
	G T	
1641	TTTCATCCATTTTTTCCAATACAGTACCAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTTGGTTATTTTG	1720
	C G C C C C C	
1721	AAAGTGCCCAATGCTTTTACATCTTCATTAGGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAAATTTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTTGAATTTATTCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
1881	GCTGTTTACGICTACAAACCAACTAGGGCTAAAAACAAAT	1920
	G C C C G C	
1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTAGTTA	1960
	G C G G	
1961	CGTATTTATCGGATGAATTTTGTCTGGATGAAAAGCGAGA	2000
	C CC CAGC G C	
2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
2041	GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTA	2080

FIGURE 10C

2081	ATAGGCAACCAGAACGTGGGTGGGGCGGAAGTACAGGGAT	2120
2121	TACCATCCAAGGAGGGGATGACGTATTTAAAGAAAATTAC G T C G C G G C	2160
2161	GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT	2200
2201	ATTTGTATCAAAAAATCGATGAATCAAAATTAAAAGCCTT CC C C G G C G C G G	2240
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA	2280
2281	GACCTAGAAATCTATTTAATTCGCTACAATGCAAAACATG C C G CC C C	2320
2321	AAACAGTAAATGTGCCAGGTACGGGTTCCCTTATGGCCGCT	2360
2361	TTCAGCCCAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
2441	GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATCATTCGCA	2480
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
2521	AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAAGA	2560
2561	CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT	2600
2601	CGAAGAGAAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
2641	AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAAAAAT G G	2680
2681	TGGAATGGGAAACAAATATCGTTTATAAAGAGGCAAAAGA G C C C C	2720
2721	ATCTGTAGATGCTTTATTTGTAAACTCTCAATATGATCAA	2760
2761	TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG	2800

FIGURE 10D

FIGURE 10E

1	ATGGATAACAAATCCGAACATCAATGAATGCATTCCTTATA	40
	C C A C A C	
41	ATTGTTTAAAGTAACCCOTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAAACTGGTTACACCCCAATCGATATTTCTCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTCTCTTTGAGTGAATTTGTTCCCGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTGTGTGTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTTGGTCCCTCTCAATGGGACGCATTCTTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTCTCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCTTACAACCGCTATTCCTCT	440
	C C T G C A C AT	
441	TTTTCAGTTCAAATATATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTGTAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTGTTGGACAAAGGTGGGGATTGTGATCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	
641	GATTAGAACGTGTATGGGACCGGATTCTAGAGATTGGGT	680
	C G G C T T A	

FIGURE 11A

681	AAGGTATAATCAATTTAGAAGAGAATTAACTAACTGTA	720
	T A C C G C G G C C A T	
721	TTAGATATCGTTGCTGCTGTCCCGAATTATGATAGTAGAA	760
	G C T GT C C CTCC	
761	GATATCCAATTTCGAACAGTTTCCCAATTAAACAAGAGAAAT	800
	CC C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAAATTTTGATGGTAGTTTT	840
	C T TC T G C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAAAGATATTAGGAGTC	880
	T T T C A T C G CTCC C C	
881	CACATTGATGGATATACCTTAACAGTATAACCATCTATAC	920
	C C CT G C T C	
921	GGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATCAA	960
	C C A AG G C T A C G	
961	ATAATGGCTTCTCCTGTAGGGTTTTTCGGGGCCAGAATTCA	1000
	C C A T A CAGC C G T	
1001	CTTTCCGCTATATGGAACATATGGGAAATGCAGCTCCACA	1040
	C T C C C	
1041	ACAACGTATTGTTGCTCAACTAGGTCAGGCGGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTTAAATATAG	1120
	C G T C G C C C C	
1121	GGATAAATAATCAACAACCTATCTGTCTCTTGACGGGACAGA	1160
	T C C C G T C A	
1161	ATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTGCGTGGATGAAATAC	1240
	G C T CT C C	
1241	CGCCACAGAAATAACAACGTGCCACCTAGGCAAGGATTATG	1280
	A C T C CTC	
1281	TCATCGATTAAAGCCATGTTTCAATGTTTCGTTTCAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	
1361	CTTGGATACATCGTAGTGCTGAATTTAATAATATAATTGC	1400
	C G C C C C C	

FIGURE 11B



1401	ATCGGATAGTATTACTCAAATCCTCGAGTGAAGGGAAC	1440
	C	
1441	TTTCTTTTAAATGGTTCTGTAATTCAGGACCAGGATTTA	1480
	C C C C	
1481	CTGGTGGGACTTAGTTAGATTAAATAGTAGTGGAAATAA	1520
	A C C C C C	
1521	CATTCAGAAATAGAGGGTATATTGAAGTTCCAATTCAC TTC	1560
1561	CCATCGACATCTACCAGATATCGAGTTCGTGTACGGGTATG	1600
	C A GA	
1601	CTTCTGTAACCCCGATTACCTCAACGTTAATTGGGGTAA	1640
	G T	
1641	TTCATCCATTTTTTTCCAATACAGTACCAGCTACGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTTGGTTATTTTG	1720
	C G C C C C	
1721	AAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTTGAATTTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGAAAGAGCGCAGAAGGCGGTGAATGC	1880
	G C C T G C T C	
1881	GCTGTTTACGTTCTACAAACCACTAGGGCTAAAAACAAAT	1920
	C C C C C T G T CT G T C	
1921	GTAACGGATTATCATATTGATCAAGTGTCOAATTTAGTTA	1960
	T T C C C C G C	
1961	CGTATTTATCGGATGAATTTTGCTGGATGAAAAGCGAGA	2000
	C CC TAGC G C C C C G T	
2001	ATTGTCGAGAAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
	C C T C C T C C	
2041	GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTA	2080
	GA G C CT G C C C C	
2081	ATAGGCAACCCAGAACGTGGGTGGGGCGGAAGTACAGGGAT	2120
	C G T T C C	

FIGURE 11C

2121	TACCATCCAGGAGGGGATGACCTATTTAAAGAAAATTAC	2150
	C C C T G C G G C	
2161	GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT	2200
	C C C A T C C C T C	
2201	ATTTGTATCAAAAAATCGATGAATCAAAATTAAAAGCCTT	2240
	C C G G G C C C	
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA	2280
	C A G C T C C C C	
2281	GACTTAGAAATCTATTTAATTCGCTACAATGCAAAACATG	2320
	C T C C G C A G C G C	
2321	AAACAGTAAATGTGCCAGGTACGGTTCCTTATGGCCGCT	2360
	G C G C T C C A	
2361	TTCAGCCCCAAAGTCCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
	T T C T G T C	
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
	A T G G C	
2441	GTTCTGTAGGGATGGAGAAAAGTGTGCCCATCATTCGCA	2480
	C C C C G C T	
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
	C G C G T C G	
2521	AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAAGA	2560
	C A C C C C	
2561	CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT	2600
	C C A T C C T	
2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
	G C T T C	
2641	AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAAAAAT	2680
	G A G G G G C	
2681	TGGAATGGGAAACAAATATCGTTTATAAAGAGGCCAAAGA	2720
	C T C C G C	
2721	ATCTGTAGATGCTTTATTTGTAACTCTCAATATGATCAA	2760
	G C G G C G C C	
2761	TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG	2800
	G C C C C C C C C	
2801	ATAAACGTGTTTCATAGCATTGGAAGCTTATCTGCCTGA	2840
	C G C T G C T	

FIGURE 11D

2841	GCTGCTCTGTGATTCCGGGTTGTCAATGCGGCTATTTTGAA	2880
	T C C T G C T C C C G	
2881	GAATTAGAAGGGCGTATTTTCACTGCATTCTCCCTATATG	2920
	C T G A C T C T G C	
2921	ATGCGAGAAATGTCATTAAAAATGGTGATTTTAATAATGG	2960
	C C C G C C C C	
2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000
	C CAG T T G C G G	
3001	GAACAAAACAACCAACGTTCCGGTCCTTGTTGTTCCGGAAT	3040
	G T G C G T G	
3041	GGGAAGCAGAAGTGTACACAAGAAGTTCGTGTCTGTCCGGG	3080
	T C G A A A	
3081	TCGTGGCTATATCCTTCGTGTACACAGCGTACAAGGAGGGA	3120
	A A C T C G C T	
3121	TATGGAGAAGGTTGCGTAACCATTCATGAGATCGAGAACA	3160
	C T G G C C	
3161	ATACAGACGAACCTGAAGTTTAGCAACTGCGTAGAAGAGGA	3200
	C C G T CTC C G A	
3201	AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT	3240
	C C C T T C C C C	
3241	GTAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA	3280
	G G G C AGC	
3281	ATCGAGGATATAACGAAGCTCCTCCGTACCAGCTGATTA	3320
	CA T C T T C	
3321	TGCGTCAGTCTATGAAGAAAATCGTATACAGATGGACGA	3360
	C C G C G G C C CA	
3361	AGAGAGAATCCTTGTTGAATTAAACAGAGGGTATAGGGATT	3400
	C T C C G C T C C	
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
	A T C T C G GC T	
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
	G T T G C A G C C T	
3481	GAAACGGAAGGAACATTTATCGTGGACAGCGTGAATTAC	3520
	C G C C GC T	
3521	TCCTTATGGAGGAA	3534
	T G	

FIGURE 11E

```

1  ATGACTGCAGATAATAATACGGAAGCACTAGATAGCTCTA  40
   C C C C C C C C T
41  CAACAAAAGATGTGCATTCAAAAAGGCATTCCGTAGTAGG  80
   C T G T C G G T C T G
81  TGATCTCCTAGGCGTAGTAGGTTTCCCGTTTGGTGGAGCG  120
   A C T G G T A T C C C
121 CTTGTTTCGTTTTATACAAACTTTTTAAATACTATTGGC  160
   C GAGC C C C C C
161 CAAGTGAAGACCCCGTGAAGGCTTTTATGGAACAAGTAGA  200
   C G T A A C G T
201 AGCATTGATGGATCAGAAAATAGCTGATTATGCAAAAAAT  240
   TC T G T A C G C
241 AAAGCTCTTGCAAGTACAGGGCCTTCAAAATAATGTGG  280
   G T G AC C G C G
281 AAGATTATGTGAGTGCATTGAGTTCATGGCAAAAAATCC  320
   G C C TCCAGC G G C
321 TGTGAGTTCACGAAATCCACATAGCCAGGGGCGGATAAGA  360
   T C CA T C A TA C
361 GAGCTGTTTTCTCAAGCAGAAAGTCATTTTCGTAATTCAA  400
   T C C TCC C CA A C
401 TGCCTTCGTTTGCAATTTCTGGATACGAGGTTCTATTTC  440
   AGC T C C T T C
441 AACACATATGCACAAGCTGCCAACACACATTTATTTTAA  480
   C T C T C C G C C
481 CTAAGAAGACGCTCAAATTTATGGAGAAGAATGGGGATACG  520
   T G C G
521 AAAAAGAAGATATTGCTGAATTTTATAAAGACAATAAA  560
   G G C G C C GC T T
561 ACTTACGCAAGAATATACTGACCATTGTGTCAAATGGTAT  600
   G C C G C C G
601 AATGTTGGATTAGATAAAATTAAGAGGTTTATCTTTATGAAT  640
   C TC C GC C C T C C G
641 CTTGGGTAAACTTTAACCGTTTATCGCAGAGAGATGACATT  680
   G C A A CA G C

```

FIGURE 12A

681	AACAGTATTAGATTTAATTGCACCTATTTCATTGTATGAT	720
	G T GC C C T C C C C	
721	GTTCCGGCTATACCCAAAAGAAGTTAAAACCGAATTAACAA	760
	GA A C G G T GC T C	
761	GAGACGTTTTAACAGATCCAATTGTCGGAGTCAACAACCT	800
	GC C T C T	
801	TAGGGGCTATGGAACAACCTTCTCTAATATAGAAAATTAT	840
	T T AGC C C C	
841	ATTCGAAAACCATCTATTTGACTATCTGCATAGAAATTC	880
	A G C C T C	
881	AATTTCACACGCGGTTCCAAACCAGGATATTATGGAAATGA	920
	C AA T C T C	
921	CTCTTTCAATTATTGGTCCGGTAATTATGTTTCACTAGA	960
	C C C C C	
961	CCAAGCATAGGATCAAATGATATAATCACATCTCCATTCT	1000
	T T C C C	
1001	ATGGAAATAAATCCAGTGAACCTGTACAAAATTAGAAATT	1040
	T C G G G CC T G	
1041	TAATGGAGAAAAAGTCTATAGACCGTAGCAAATACAAAT	1080
	C C C G C C C	
1081	CTTGCGGTCTGGCCGTCGCTGTATATTAGGTGTTACAA	1120
	C T G A A T C C C	
1121	AAGTGAATTAGCCAAATATAATGATCAACAGATGAAGC	1160
	G G T G C G C G	
1161	AAGTACACAAACGTACGACTCAAAAAGAAATGTTGGCGCG	1200
	C C C G T C C T C A	
1201	GTCAGCTGGGATTCTATCGATCAATTGCCTCCAGAAACAA	1240
	TCT C C	
1241	CAGATGAACCTCTAGAAAAGGGATATAGCCATCAACTCAA	1280
	C AT G G C C C T	
1281	TTATGTAATGTGCTTTTTAATGCAGGGTAGTAGGAACA	1320
	C G C G A TCC G C	
1321	ATCCAGTGTAACTTGGACACATAAAAGTGTAGACTTTT	1360
	T G C C GTCC G C	
1361	TTAACATGATGATTGCAAAAAAATTACACAACCTCCGTT	1400
	C AGC G G C T C	

FIGURE 12B

1401	AGTAAAGGCATATAAGTTACAATCTGGTGCTCCGTTGTC	1440
	G G A C C C G	
1441	GCAGGTCCTAGGTTTACAGGAGGAGATATCATTCAATGCA	1480
	C A C T T C C G	
1481	CAGAAAATGGAAGTGCGGCAACTATTTACGTTACACCGGA	1520
	G C C C A T C G T	
1521	TGTGTCGTA CTCTCAAAAATATCGAGCTAGAATTCATTAT	1560
	T G G C A G A C T C	
1561	GCTTCTACATCTCAGATAACATTTACACTCAGTTTAGACG	1600
	A CAGC C C C C G T	
1601	GGGCACCATTTAATCAATACTATTTTCGATAAAACGATAAA	1640
	A C C C G T C T C G C C	
1641	TAAAGGAGACACATTAAACGTATAATTCATTTAATTTAGCA	1680
	C T TC C A C A G C C C G	
1681	AGTTTCAGCACACCATTTCGAATTATCAGGGAATAACTTAC	1720
	T C C C C TC T	
1721	AAATAGGCGTCACAGGATTAAGTGCTGGAGATAAAGTTTA	1760
	G C C TC C C C C C	
1761	TATAGACAAAATTGAATTTATTCCAGTGAAT	1791
	C C G G C C C	

FIGURE 12C

1	ATG AATAATGTATTGAATAGTGGGAAGAACAACTATT	40
	GAC C C C CTC T C C	
41	GTGATGCGTATAATGTAGTAGCCCATGATCCATTAGITT	80
	C C A C C C G T C C C	
81	TGAACATAAAATCATTAGATACCATCCAAAAGAATGGATG	120
	C C GAGCC C C T T G G G	
121	GAGTGGAAAAGAACAGATCATAGTTTATATGTAGCTCCTG	160
	A C T T C CTC C C C A	
161	TAGTCGGAACGTGTCTAGTTTTTTGCTAAAGAAAGTGGG	200
	G T A C C CC T C G C	
201	GAGTCTTATTGGAAAAAGGATATTGAGTGAATTATGGGGG	240
	CTC C C C T C TCC C C T	
241	ATAATATTTCTAGTGGTAGTACAAATCTAATGCAAGATA	280
	C C ATC GTCC T C C	
281	TTTTAAGGGAGACAGAACAATTCCTAAATCAAAGACTTAA	320
	C G C G T C C GC T C	
321	TACAGATACCCTTGCTCGTGTAATGCAGAATTGATAGGG	360
	C T T G A A C C T G C T	
361	CTCCAAGCGAATATAAGGGAGTTTAAATCAACAAGTAGATA	400
	A C TC T C C G G C	
401	ATTTTTTAAACCCTACTCAAACCCTGTTCCTTTATCAAT	440
	C C G T A G T G C T C	
441	AACCTCTTCGGTTAATACAATGCAGCAATTATTCTAAAT	480
	C C G C T C C C C C	
481	AGATTACCCAGTTCCAGATACAAGGATACCAGTTGTTAT	520
	G T T T C C C	
521	TATTACCTTTATTTGCACAGGCAGCCAATATGCATCTTTC	560
	TC T AC C T T C CT G	
561	TTTTATTAGAGATGTTATTCCTAATGCAGATGAATGGGGT	600
	C C AC T C G C C C T C A	
601	ATTCAGCAGCAACATTACGTACGTATCGAGATTACCTGA	640
	C T C TC TA G A CA C T	
641	GAAATTATACAAGAGATTATTCCTAATTATTGTATAAATAC	680
	G C C TC T C C C C C C	

FIGURE 13A

681	STATCAAACCTGCGTTTAGAGGGTTAAACACCCGTTTACAC	720
	T G C C T A C C T T A G C T	
721	GATATGTTAGAAATTTAGAACATATATGTTTTTAAATGTAT	760
	C C T G C G C C C C T C G	
761	TTGAATATGTATCCATTGGTCAATTGTTTAAATATCAGAG	800
	G C CAG AGTC C C G C	
801	TCTTATGGTATCTCTCTGGCGCTAATTTATATGCTAGCGGT	840
	CT G G C A C C C C CTCT C	
841	AGTGGACCACAGCAGACACAATCATTACAGCACAAAAC	880
	A T GAGC C T G	
881	GGCCATTTTTATATTCTCTTTTCCAAGTTAATTCGAATTA	920
	C G AGCT G C C C C	
921	TATATTATCTGGTATTAGTGGTACTAGGCTTCTATTACC	960
	C TC CAG CTC G C A C C A	
961	TTCCCTAATATTGGTGGTTTACCGGGTAGTACTACAAC	1000
	T C C AC T A CTCC C	
1001	ATTCATTGAATAGTGCCAGGGTTAATTATAGCGGAGGAGT	1040
	AGCC T CTC A G C C T T	
1041	TTCATCTGGTCTCATAGGGGCGACTAATCTCAATCACAAC	1080
	CAGC AT G T T A CT G C	
1081	TTTAATTGCAGCACGCTCCTCCCTCTTTATCAACACCAT	1120
	C TC C T G A C GAGC G	
1121	TTGTTAGAAAGTTGGCTGGATTACAGTACAGATCGAGAGGG	1160
	G GTCC T CAGC T C A	
1161	CGTTGCTACCTCTACGAATTGGCAGACAGAATCCTTTCAA	1200
	A A C A C G C	
1201	ACAACTTTAAGTTTAAAGGTGTGGTGCTTTTTCAGCCCGTG	1240
	C C T CC TC A C T A	
1241	GAAATTCAAAACATTTTCCCAGATTATTTTATCCGTAATAT	1280
	G C T C C C TA G C	
1281	TTCTGGGGTTCCTTTAGTTATTAGAAACGAAGATCTAACA	1320
	C T C C C C G T C C C	
1321	AGACCGTTACACTATAACCAAATAAGAAATATAGAAAGTC	1360
	C T AC T T C G T G C C GTC	
1361	CTTCGGGAACACCTGGTGGAGCACGGGCCTATTGGGTATC	1400
	A C T T A A T A A T CC C G	

FIGURE 13B



1401	TGTGCATAACAGAAAAAATAATATCTATG <sup>˙</sup> CCGCTAATGAA <sup>˙</sup>	1440
	C G G C C C T C C G	
1441	AATGGTACTATGATCCATT <sup>˙</sup> TGGCGCCAGAAGATTATACAG <sup>˙</sup>	1480
	C C T CC T A C T	
1481	GATTTACTATATCGCCAATACATGCCACTCAAGTGAATAA <sup>˙</sup>	1520
	C C C T C T C C	
1521	TCAAAC <sup>˙</sup> TCGAACATTTATT <sup>˙</sup> TCTGAAAAATTTGGAATCAA <sup>˙</sup>	1560
	G A C C C C C G C	
1561	GGTGATTCCTTAAGATT <sup>˙</sup> TGAACAAAGCAACACGACAGCTC <sup>˙</sup>	1600
	C G G C G TC T C A	
1601	GTTATACGCTTAGAGGGAATGGAAATAGTTACAATCTTTA <sup>˙</sup>	1640
	G C TT G C C C	
1641	TTTAAGAGTATCTTCAATAGGAAATCAACTATT <sup>˙</sup> CGAGTT <sup>˙</sup>	1680
	C G TAGC C T T C C C C T	
1681	ACTATAAACGGTAGAGTTTATACTGTTTCAAATGTTAATA <sup>˙</sup>	1720
	C C AC T C A C T G C	
1721	CCACTACAAATAACGATGGAGTTAATGATAATGGAGCTCG <sup>˙</sup>	1760
	T A G C T C C C CA	
1761	TTTTTCAGATATTAATATCGGTAATATAGTAGCAAGTGAT <sup>˙</sup>	1800
	A CAGC C C C T C C C G CTC C	
1801	AATACTAATGTAACGCTAGATATAAATGTGACATTAAACT <sup>˙</sup>	1840
	C C T TT G C C CC C T	
1841	CCGGTACTCCATT <sup>˙</sup> TGATCTCATGAATATTATGTTTGTGCC <sup>˙</sup>	1880
	T A C C	
1881	AACTAATCTTCCACCACTTTAT <sup>˙</sup>	1902
	C C T T G C	

FIGURE 13C

1	ATGGAGGAAATAATCAAATCAATGCATACCTTACAATT	40
	G C C C T A C	
41	GTTTAAAGTAATCCTGAAGAAGTACTTTTGGATGGAGAACG	80
	C G C A G T GC T	
81	GATATCAACTGGTAATTCATCAATTGATATTTCTCTGTCA	120
	C T C C T C C C CT C	
121	CTTGTTTCAGTTTCTGGTATCTAAGTTTGTACCAGGGGGAG	160
	T G C CAGC C G T T	
161	GATTTTGTAGTTGGATTAATAGATTTTGTATGGGGAATAGT	200
	G CC T C C T C C C T C	
201	TGGCCCTTCTCAATGGGATGCATTCTAGTACAAATTGAA	240
	T A C G G G	
241	CAATTAAATTAATGAAAGAATAGCTGAATTTGCTAGGAATG	280
	G G C C G G C G C C C	
281	CTGCTATTGCTAATTTAGAAGGATTAGGAAACAATTCAA	320
	C C C G G C T C	
321	TATATATGTGGAAGCATTAAAGAATGGGAAGAAGATCCT	360
	C C G C C G C	
361	AATAATCCAGAAACCAGGACCAGAGTAATTGATCGGTTTC	400
	C G C C T G G C CA A CA	
401	GTATACTTGATGGGCTACTTGAAAGGGACATTCCTTCGTT	440
	A CT G C C CT G G A T C A C	
441	TCGAATTTCTGGATTGGAAGTACCCCTTTATCCGTTTAT	480
	CA C C C T T C G G C	
481	GCTCAAGCGGCCAATCTGCATCTAGCTATATTAAGAGATT	520
	A T T C C CC TC CA	
521	CTGTAAATTTTGGAGAAAGATGGGGATTGACACGATATAA	560
	G C C G G C T C	
561	TGTCAATGAAAACATAATAGACTAATTAGGCATATTGAT	600
	C G T C C T C C	
601	GAATATGCTGATCACTGTGCAATACGTATAATCGGGGAT	640
	G C C C T C C C C T C	
641	TAAATAATTTACCGAAATCTACGTATCAAGATTGGATAAC	680
	G C C C T G T T	
681	ATATAATCGATTACGGAGAGACTTAACATTGACTGTATTA	720
	C C CA G GA G CC C A T G	

FIGURE 14A

721	GATATCGCCGCTTTCTTTCCAACTATGACAATAGGAGAT	760
	C T A C G C	
761	ATCCAATTCAGCCAGTTGGTCAACTACAAAGGGAAGTTTA	800
	C T C A G T C A C	
801	TACGGACCCATTAATTAATTTTAATCCACAGTTACAGTCT	840
	T C T C C C T G AAG	
841	GTAGCTCAATTACCTACTTTTAACGTTATGGAGAGCAGCC	880
	C C C T C A C C TC	
881	GAATTAGAAATCCTCATTATTTGATATATTGAATAATCT	920
	T C G C A C G C C C C C	
921	TACAATCTTTACGGATTGGTTTAGTGTGGACGCAATTTT	960
	T C C C C C G T C C	
961	TATTGGGGAGGACATCGAGTAATATCTAGCCTTATAGGAG	1000
	T C A G C C CTCT T	
1001	GTGGTAACATAACATCTCCTATATATGGAAGAGAGGCGAA	1040
	G T C C C T A	
1041	CCAGGAGCCCTCCAAGATCCTTTACTTTTAATGGACCGGTA	1080
	A C TAGT G C C C T A C	
1081	TTTAGGACTTTATCAAATCCTACTTTACGATTATTACAGC	1120
	C A C G T C C GA GC C.	
1121	AACCTTGGCCAGCGCCACCATTTAATTTACGTGGTGTGA	1160
	T T C CC TA A	
1161	AGGAGTAGAATTTTCTACACCTACAAATAGCTTTACGTAT	1200
	G C T G C T C CTC C T C	
1201	CGAGGAAGAGGTACGTTGATTCTTTAACTGAATTACCGC	1240
	A T A C C G C C C A	
1241	CTGAGGATAATAGTGTGCCACCTCGCGAAGGATATAGTCA	1280
	A C C CA G C CTCC	
1281	TCGTTTATGTCATGCAACTTTTGTTCAAAGATCTGGAACA	1320
	CA G G C C C C G GC T C T	
1321	CCTTTTTTAAACAAGTGGTGTAGTATTTTCTTTGGACCGATC	1360
	A CC C T A A T G C A T	
1361	GTAGTGCAACTCTTACAAATACAATTGATCCAGAGAGAAT	1400
	T C T C C G	

FIGURE 14B

1401	TAATCAAATACCTTTAGTGAAGGATTTAGAGTTTGGGGG	1440
	C A G C G T C C T G A	
1441	GGCACCTCTGTCATTACAGGACCAGGATTTACAGGAGGGG	1480
	A T C C C T	
1481	ATATCCTTCGAAGAAATACCTTTGGTGATTTTGTATCTCT	1520
	T A C T C C GAGC	
1521	ACAAGTCAATATTAATTCACCAATTACCCAAAGATACCGT	1560
	C T C C C T T T	
1561	TTAAGATTTGCTTACGCTTCCAGTAGGGATGCACGAGTTA	1600
	C C G A TTCC T C TA C	
1601	TAGTATTAACAGGAGCGGCATCCACAGGAGTGGGAGGCCA	1640
	C GC C C C A T T C T C T A	
1641	AGTTAGTGTAATATGCCTCTTCAGAAAATATGGAATA	1680
	CTCC G C A C G G C	
1681	GGGGAGAACTTAACATCTAGAACATTTAGATATACCGATT	1720
	C G C G C C C C	
1721	TTAGTAATCCTTTTTCATTTAGAGCTAATCCAGATATAAT	1760
	CTC C CAGT CC T C C T C C	
1761	TGGGATAAGTGAACAACCTCTATTGGTGACGGTTCTATT	1800
	C T C C A T AGC C	
1801	AGTAGCGGTGAACCTTTATATAGATAAAATTGAAATTATTC	1840
	TCATCT C T G C T C G G C	
1841	TAGCAGATGCAACATTTGAAGCAGAATCTGATTTAGAAAAG	1880
	T C C T CC C G T G ACA CC T G	
1881	AGCACAAAAGCGGTGAATGCCCTGTTTACTTCTTCCAAT	1920
	C G T C C C CA	
1921	CAATCGGGTTAAAAACCGATGTGACGGATTATCATATTG	1960
	GC T C G TA C T T C C	
1961	ATCAAGTATCCAATTTAGTGGATTGTTTATCAGATGAATT	2000
	C G C G CACC ACC TAGC G	
2001	TTGTCTGGATGAAAAGCGAGAATTGTCCGAGAAAGTCAAA	2040
	C C C C G T C C T	
2041	CATGCGAAGCGACTCAGTGATGAGCGGAATTTACTTCAAG	2080
	C C T C C A C CT G	
2081	ATCCAAACTTCAGAGGGATCAATAGACAACCGACCGGTGG	2120
	CT C A AC C G G A	

FIGURE 14C

2121	CTGGAGAGGGAAGTACAGATATTACCATCCAAGGAGGAGAT	2160
	T G T C C G G C C C	
2161	GACGTATTCAAAGAGAATTACGTACACTACCGGGTACCG	2200
	T G G C C C T C A T T	
2201	TTGATGAGTGCTATCCAACGTATTATATCAGAAAAAG	2240
	C C C T C C G C G C	
2241	TGAGTCGAAATTAAAAGCTTATACCGTTATGAATTAAGA	2280
	C C C C T C A G C C T	
2281	GGGTATATCGAAGATAGTCAAGACTTAGAAATCTATTGA	2320
	C C C C C T C C	
2321	TCCGTTACAATGCAAAACACGAAATAGTAAATGTGCCAGG	2360
	A G C G G C C G C C	
2361	CACGGGTTCCCTTATGGCCGCTTTCAGCCCCAAATGCCAATC	2400
	T T C C A T T C T C T	
2401	GGAAAGTGTGGAGAACCGAATCGATGCGCGCCACACCTTG	2440
	G G T C A T	
2441	AATGGAAATCCTGATCTAGATTGTTCTCTGCAGAGACGGGGA	2480
	G C T G C C G T C	
2481	AAAATGTGCACATCATTCCCATCATTTACGCTTGGATATT	2520
	G G C C T C T C C	
2521	GATGTTGGATGTACAGACTTAAATGAGGACTTAGGTGTAT	2560
	G T C G C C A C	
2561	GGGTGATATTCAAGATTAAAGACGCAAGATGGCCATGCAAG	2600
	C C C C C A C	
2601	ACTAGGGAATCTAGAGTTTCTCGAAGAGAAACCATTATTA	2640
	T C C T G G C	
2641	GGGGAAGCACTAGCTCGTGTGAAAAGAGCGGAGAAGAAGT	2680
	T T C G A	
2681	GGAGAGACAAACGAGAGAACTGCAGTTGGAACAAATAT	2720
	G T C G A G T C	
2721	TGTTTATAAAGAGGCAAAAGAATCTGTAGATGCTTTATTT	2760
	C C G C G G C	
2761	GTAACTCTCAATATGATAGATTACAAGTGGATACGAACA	2800
	G C C A G G C C C C	
2801	TCGCCATGATTTCATGCGGCAGATAAACGCGTTCATAGAAAT	2840
	C C C C T G C C	

FIGURE 14D

2841 CCGGGAAGCGTATCTGCCAGAGTTGTCTGTGATTCCAGGT 2880  
 T T G T CT T C C T  
 2881 GTCAATGCGGCCATTTTCGAAGAATTAGAGGGACGTATTT 2920  
 G C T C G C T C  
 2921 TTACAGCGTATTCTTATATGATGCGAGAAATGTCATTAA 2960  
 C A TC G C C C C  
 2961 AAATGGCGATTTCATAATGGCTTATTATGCTGGAACGTG 3000  
 G C T C C C CAGC T  
 3001 AAAGGTCATGTAGATGTAGAAGAGCAAAACAACCACCGTT 3040  
 G C G G A G T G  
 3041 CGGTCCTTGTATCCAGAATGGGAGGCAGAAGTGTACAA 3080  
 C G G G T G A T C  
 3081 AGAGGTTGCTGTCTGTCCAGGTCGTGGCTATATCCTTCGT 3120  
 A A A A C T C  
 3121 GTCACAGCATATAAAGAGGGATATGGAGAGGGCTGCGTAA 3160  
 G C T C G C T T G  
 3161 CGATCCATGAGATCGAAGACAATACAGACGAACTGAAATT 3200  
 C C C GA C C G T G  
 3201 CAGCAACTGTGTAGAAGAGGAAGTATATCCAAACAACACA 3240  
 TC C C G A A C C  
 3241 GTAACGTGTAATAATTATCTGGGACTCAAGAAGAATATG 3280  
 T T C CG C C T A G G C  
 3281 AGGGTACGTACACTTCTCGTAATCAAGGATATGACGAAGC 3320  
 GA G C AGC CAG T CA  
 3321 CTATGGTAATAACCCCTTCGTTACCAGCTGATTACGCTTCA 3360  
 TCC TCXXXXXXX T T C T C C  
 3361 GTCTATGAAGAAAAATCGTATACAGATGGACGAAGAGAGA 3400  
 G C G G C C CA C T  
 3401 ATCCTTGTGAATCTAACAGAGGCTATGGGGATTACACACC 3440  
 C C G TC T CA C  
 3441 ACTACCGGCTGGTTATGTAACAAGGATTTAGAGTACTTC 3480  
 T A T C T C GC T T  
 3481 CCAGAGACCGATAAGGTATGGATTGAGATCGGAGAAACAG 3520  
 T C A G C T C  
 3521 AAGGAACATTCATCGTGGATAGCGTGGAAATCTCCTTAT 3560  
 G C C GC T T G  
 3561 GGAGGAA 3567

FIGURE 14E

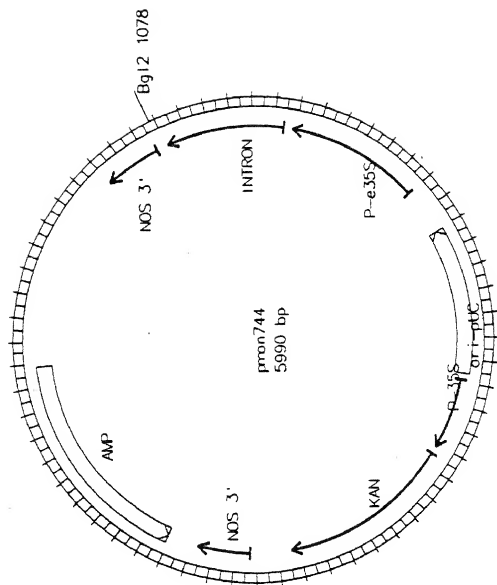


FIGURE 15

1	AGATCTAGAGGTAATTTGTTATGAGTACTGTCGTTGGTTAAG	40
	GATC	
41	GGAAACGTCAACGGTGGTGTACAACAACCTAGAAGGAGGA	80
	G T A	
81	GAAGGCAATCCCTTCGCAGGAGGGCTAACAGAGTACAGCC	120
	T A T	
121	AGTGGTTATGGTCACTGCTCCTGGCGAACCAGGAGGAGG	160
	GC A A A	
161	AGACGCAGAAGAGGAGGCAATCGCAGGTCAAGAAGAACTG	200
	A G T A	
201	GAGTTCCAGGGGAAGGGGCTCAAGCGAGACATTCTGTGT	240
	A A T	
241	TACAAAGGACAACCTCGTGGGCAACTCCCAAGGAAGTTTC	280
281	ACCTTCGGACCAAGTGTATCAGACTGTCCAGCATTCAAGG	320
	T	
321	ATGGAATACTCAAGGCCTACCATGAGTACAAGATCACAAG	360
	T	
361	TATCCTTCTTCAGTTCGTCAGCGAGGCCTCTTCCACCTCA	400
	T G T	
401	CCAGGATCCATCGCTTATGAGTTGGACCCACATTGCAAAG	440
	C A T	
441	TATCATCCCTCCAGTCCTACGTCAACAAGTTCCAAATCAC	480
	T	
481	AAAGGGAGGAGCTAAGACCTATCAAGCTAGGATGATCAAC	520
	T T C T	
521	GGAGTAGAATGGCACGATTCTCTGAGGATCAGTGCAGGA	560
	T T A	
561	TACTTTGGAAGGAAGTGGAAAATCTTCAGACCCAGCAGG	600
	C A G T T	
601	ATCTTTCAGAGTCACCATCAGAGTGGCTCTTCAAAACCCC	640
	T T A	
641	AAGTAATAGACTCCGGATCAGAGCCTGGTCCAAGCCCACA	680
	A T	

FIGURE 16A



```

681  ACCAACACCCACTCCAAC·TCCCCAAAGCATGAGCGATT· 720
721  ATTGCTTACGTCGGCATACCTATGCTGACCATTCAAGAAT· 760
761  TC 762

```

FIGURE 16B